Developmental Reading and English Assessment in a Researcher-Practitioner Partnership

Dolores Perin
Julia Raufman
Hoori Santikian Kalamkarian

December 2015

CCRC Working Paper No. 85

Address correspondence to:

Dolores Perin
Professor of Psychology and Education and
Senior Research Associate, Community College Research Center
Teachers College, Columbia University
525 W 120th St, Box 174
New York, NY 10027
Email: perin@tc.edu

This research was funded by the Bill & Melinda Gates Foundation under the project entitled "Analysis of Statewide Developmental Education Reform: Learning Assessment Study." The authors would like to thank the other members of the assessment team, Nikki Edgecombe, Susan Bickerstaff, and Madeline Joy Trimble, for their suggestions for the study and review of earlier drafts of this paper. Mark Lauterbach provided helpful assistance with the statistical analyses. We also appreciate the assistance of Rachel Kasten, Kristen Kurre, Mollie Book, Korleen Brady, and Carolyn Foster in scoring and data entry, and of Geremy Grant and Allyson Shaw in the analysis of student interviews.

Abstract

This paper reports findings from a researcher-practitioner partnership that assessed the readiness for postsecondary reading and writing demands of 211 students in developmental reading and English courses in two community colleges. An assessment battery was designed for the study, comprising two standardized tests and five project-developed tasks. The project-developed measures were two text-based writing tasks similar to those typically assigned in college classrooms (a summarization task and a persuasive essay), a self-efficacy scale, a teacher judgment questionnaire, and a qualitative student retrospective report. The text-based writing measures were keyed to high-enrollment, introductory-level general education courses that had significant literacy demands.

The results pointed to areas where students needed improvement in order to be ready for literacy tasks at the introductory postsecondary level. There was a discrepancy between the relatively low reading and writing skills as assessed through performance tasks and relatively high student self-efficacy ratings and teacher judgments. This finding suggests the possibility of an unrealistic amount of confidence in students' ability to perform college-level reading and writing tasks. Correlations between assessment measures tended to be moderate, suggesting that the measures were tapping different skills. A series of hierarchical regressions modeling the text-based writing skills suggested that improvement in text-based summarization may require particular attention to reading comprehension skills, while improvement in text-based persuasive essay writing may depend more on developing general writing skills. Students' retrospective reports indicated that although participants had some difficulty stating the requirements of the summarization task, they described appropriate strategies to complete it. Overall, the study's findings point to the need to examine approaches to instruction, curriculum, course structure, and placement policy that may improve students' college readiness.

Table of Contents

1. Introduction

Despite completing secondary education, many students in the United States enter postsecondary institutions with low reading, writing, and/or mathematics skills, which greatly impedes their academic progress (J. Jackson & Kurlaender, 2014; Porter & Polikoff, 2012; Sparks & Malkus, 2013). In fall 2000, 42 percent of entering students at two-year public colleges and 20 percent of entering students at four-year public colleges enrolled in at least one developmental education course (Parsad & Lewis, 2003)[1] aimed at preparing students for the academic demands of postsecondary coursework. The current study focuses on reading and writing ability, which prior research indicates is problematic for a large proportion of postsecondary students. For example, in a sample of 57 community colleges in seven states, 33 percent of entering students were referred to developmental reading courses (Bailey, Jeong, & Cho, 2010). Information on the need for developmental English[2] courses is limited, but studies suggest that up to 35 percent of entering community college students are referred to such courses (Jenkins & Boswell, 2002; Perin & Charron, 2006). Thus, it is probable that at least one third of entering community college students require help with reading and/or writing skills if they are going to perform well in college-level courses.

Although academic skills are not the only measure of college readiness, they are a central indicator (Armstrong, Stahl, & Kantner, 2015; National Center on Education and the Economy, 2013). Signs of college readiness include passing scores on reading, writing, and mathematics placement tests administered on entry to college, and passing grades in entry-level, college-credit English composition courses (Lym, 2014). More generally, a well-prepared secondary education graduate has been characterized as one who "can qualify for and succeed in entry-level, credit-bearing college courses leading to a baccalaureate or certificate, or career pathway-oriented training programs without the need for remedial or developmental coursework" (Conley, 2012, p. 1).

From the perspective of literacy, college readiness includes the ability to read analytically and critically, synthesize written information, and produce ideas in writing

[1] Student-reported data lower enrollment rates to 24 percent at publicly funded community colleges (Sparks & Malkus, 2013). This estimate should be regarded cautiously because it is based on self-report rather than on institutional data.

[2] Developmental English courses teach basic writing skills.

that are well supported, well organized, and expressed using appropriate grammar and academic style. Well-prepared students are able to decipher the majority of words in the text they are expected to read, adapt their reading and writing skills to suit different purposes and audiences, and assess and reflect on their own skills (Atkinson, Zhang, Phillips, & Zeller, 2014; Conley, 2008; Fallahi, 2012; Mongillo & Wilder, 2012; Theurer, 2011; Wang, 2009; Yancey, 2009). However, although several general indicators of academic readiness for college exist, there is a shortage of research on how close developmental education students are to being able to perform reading and writing tasks typically required in college-credit courses.

This paper reports the results of a researcher-practitioner partnership project that assessed the readiness of developmental reading and English students for the literacy demands of introductory college-credit courses in key content areas. In this work, we went beyond conventional assessment practices. Whereas traditional methods utilize standardized tests that are not designed to reflect college-level curriculum demands, our approach includes, in addition to standardized tests, tasks intended to capture students' ability to perform typical reading and writing tasks at the introductory college level, as well as self-efficacy ratings, teacher judgments, and retrospective reports that provide first-person insight into performance.

We begin this paper by discussing a framework for college readiness in order to place the literacy assessment in context. We then discuss the literacy constructs assessed, the development of the researcher-practitioner partnership, our assessment methods, and our findings. The paper ends with a discussion of the meaning and practical implications of the results, and recommendations for future development of researcher-practitioner partnerships.

2. College Readiness

Since underprepared students experience numerous obstacles to academic achievement (Bettinger, Boatman, & Long, 2013; Cohen & Brawer, 2008), it is important to place academic skills assessment in a broader context of college readiness (Yancey, 2009). One recent and influential model of college readiness, proposed by Conley (2007),

is discussed here in order to place the study in context. This model comprises four interacting components that are proposed to affect students' ability to learn well in a postsecondary setting. Each component consists in turn of multiple subcomponents, many of which have been recognized as important by college English instructors (O'Neill, Adler-Kassner, Fleischer, & Hall, 2012).

At the heart of Conley's (2007) model is the "key cognitive strategies" component, or the work habits that support student learning. Key cognitive strategies include intellectual curiosity, an interest in inquiry, the ability to analyze and synthesize information, an understanding of the level of precision and accuracy needed to perform academic tasks, and the ability to solve problems.

A second component in the model is "key content," which covers academic content knowledge and basic reading, writing, and math skills. Conley (2007) identifies writing skill as being of central importance to this component of college readiness, especially because writing forms the basis of many assessments of knowledge in postsecondary courses: "Expository, descriptive, and persuasive writing are particularly important types of writing in college. Students are expected to write a lot in college and to do so in relatively short periods of time," and the writing should display competent grammar, spelling, and use of language (p. 14). Besides writing ability, skills in research, reading comprehension, and math, as well as disciplinary content knowledge, feature in this second component.

Interacting with key cognitive strategies and key content is a third component, "academic behaviors," which signifies students' ability to reflect on, monitor, and control their own performance. Also called metacognition (Nash-Ditzel, 2010), this component of college readiness covers understanding one's own level of mastery of a skill—for example, through assessing one's self-efficacy (Conley & French, 2014; Liao, Edlin, & Ferdenzi, 2014)—willingness to persist in difficult tasks, and an understanding of how to transfer skills to a new context.

The final component in Conley's (2007) college readiness model is "contextual skills and awareness," or a student's knowledge of the nature of college as an institution. This includes understanding academic norms and expectations as well as specific knowledge, such as of admissions, placement testing, and financial aid procedures.

The effectiveness of developmental education in promoting college readiness has been questioned in recent research (Hodara & Jaggars, 2014; Martorell & McFarlin, 2011; Melguizo, Bos, & Prather, 2011). Given the broad range of skills and behaviors required for college readiness, as detailed by Conley's (2007) model, and the multiple social and educational needs of low-achieving students, it is difficult to pinpoint the causes of this problem. However, inadequacies in assessment methods used for course placement, course structure (including multicourse sequences requiring lengthy participation), and instructional approaches have been identified as contributing to low achievement rates (Grubb & Gabriner, 2012; Hughes & Scott-Clayton, 2011). If instructional improvements are to contribute to the effectiveness of developmental education, assessment methods will be of critical importance. Assessment and instruction are intertwined, and the design of effective instruction depends on detailed knowledge of students' academic skills (Salvia, Ysseldyke, & Bolt, 2013). The current study focuses on Conley's (2007) second component, key content, in its interest in assessing developmental education students' reading and writing skills in order to gauge their level of college readiness.

3. Assessing Students' Readiness for College Literacy Demands

In this section, we describe traditional methods of assessing students' college readiness for college-level reading and writing. We then provide rationale and discuss prior research related to the constructs we measured in the current assessment. As mentioned earlier, our study expands on traditional methods by using literacy tasks that are more authentic than those used in placement tests. Another way in which the current study expands knowledge of college readiness is to include self-efficacy ratings, teacher judgments, and retrospective reports.

3.1 Traditional Methods of Assessing College Readiness

The level of college readiness in a student population has typically been measured in three ways. The first is to count college developmental education referrals and/or enrollments. In community colleges, the main site of developmental education in the

United States, up to two thirds of entering students are referred to such courses (Bailey et al., 2010), with about one third referred to developmental reading or English (Jenkins & Boswell, 2002; Perin & Charron, 2006). The second measurement method is to use cut scores established by national testing of students exiting secondary education. Using a score predictive of the ability to pass an introductory college-credit English course, ACT, a major testing company, found that 56 percent of secondary education graduates were not ready for college reading, 36 percent were not ready for college writing, and 57 percent were not ready for college math (ACT, 2014). Also based on test scores, the National Assessment of Educational Progress found that 62 percent of 12th graders were not proficient in reading, 75 percent were not proficient in writing, and 74 percent were not proficient in math (National Center for Education Statistics, 2012, 2014).

A third approach to measuring college readiness is to assess students' ability to meet specific curriculum-based reading, writing, and math standards. Reading and writing skills that are important for entering community college students have been catalogued in two studies (Armstrong et al., 2015; National Center on Education and the Economy, 2013), and a comprehensive statement of academic standards for exiting secondary education students can be found in the Common Core State Standards for College and Career Readiness (for literacy standards, see National Governors' Association Center for Best Practices & Council of Chief State School Officers, 2010). Although instruments by which college students could be assessed against such standards are not yet available, there has been interest in employing the Common Core standards in both adult basic literacy programs and postsecondary education (Holschuh, 2014; Pimentel, 2013). In the latter setting, standards currently take the form of student learning outcomes, which may be set by single institutions, college districts, or whole states (Leist, Woolwine, & Bays, 2012; Nunley, Bers, & Manning, 2011). A recent survey found that 69 percent of responding colleges used assessment rubrics representing student learning outcomes to evaluate performance on course assignments (Kuh, Jankowski, Ikenberry, & Kinzie, 2014). In the current study, we assessed students on tasks representing selected statewide learning outcomes for developmental reading and writing.

Despite many statements in the literature concerning the literacy skills entering college students should possess, there is no commonly agreed-upon empirical benchmark

in this domain (Armstrong et al., 2015). Although readiness for college literacy demands is typically operationalized in terms of placement test scores, the instruments used vary in constructs measured, and, overall, their predictive validity has been questioned (Hughes & Scott-Clayton, 2011). Thus, despite major concern in the United States over the lack of college readiness, there is no commonly agreed-upon definition of this construct that would be specific enough to translate to quantitative measures. At the same time, college instructors routinely form strong clinical judgments on whether the students in their classrooms can understand and apply material covered in their curricula (Perin & Charron, 2006). Besides teacher judgments, a conceptual framework for determining readiness for college reading and writing emerges from three sources of information: examination of text and writing assignments presented at the introductory college level (Holschuh & Aultman, 2009; MacArthur, Philippakos, & Ianetta, 2015); state, district, and/or college student learning goals for reading and writing (Barnett et al., 2012); and the Common Core State Standards for reading and writing (National Governors' Association Center for Best Practices & Council of Chief State School Officers, 2010).

3.2 Constructs Used in the Current Study

College reading and writing. Given the generally poor alignment of secondary and postsecondary literacy demands (Acker & Halasek, 2008; Williamson, 2008), it is important to know what incoming college students are in fact able to read and write. College reading requires the use of complex cognitive processes, such as analyzing text to identify the most important information, utilizing background knowledge from specific content areas, interpreting language and vocabulary appropriately for the context, consciously using personal strategies for understanding new concepts, and drawing analogies between different pieces of information (Holschuh & Aultman, 2009; Macaruso & Shankweiler, 2010; Paulson, 2014; Wang, 2009). The ability to comprehend expository, (i.e., informational), text is particularly important (Armstrong et al., 2015). At this level, competent reading depends on self-regulatory and metacognitive mechanisms, including the ability to set goals for reading a particular text, apply knowledge of text structure to the task of comprehension, assess one's understanding of information during the process of reading, and assess the trustworthiness of a particular text (Bohn-Gettler & Kendeou, 2014; Bråten, Strømsø, & Britt, 2009).

As with reading, college writing involves the use of strategies to ensure an appropriate response to an assigned prompt (Fallahi, 2012; MacArthur, Philippakos, & Ianetta, 2015). At the postsecondary level, the student writer is expected to understand the informational needs of a reader, and generate text appropriate to the purpose. At this level, students are expected to write discipline-specific texts that summarize, synthesize, analyze, and respond to information, and to offer evidence for a stated position (O'Neill et al., 2012). Further, it is expected that material be written in students' own words, and that students provide citations for quotations in order to avoid plagiarism (Keck, 2014).

Although reading and writing are often taught as separate subjects in postsecondary developmental education courses, in practice the two skills are closely related (Fitzgerald & Shanahan, 2000). In higher education, students need to integrate these skills; for example, at this level, writing assignments tend to be text-based (Carson, Chase, Gibson, & Hargrove, 1992; J. M. Jackson, 2009; McAlexander, 2003) and require critical reading of source text as the basis of a writing assignment (O'Neill et al., 2012; Yancey, 2009). The colleges in which the current research was conducted had integrated reading and writing instruction in single courses as part of statewide reform of developmental education.

Two important types of college writing are persuasive writing and written summarization (Bridgeman & Carlson, 1984; Hale et al., 1996; Wolfe, 2011). A persuasive essay requires the writer to state and defend an opinion on an issue, and, at an advanced level, to acknowledge and rebut an opposing position (De La Paz, Ferretti, Wissinger, Yee, & MacArthur, 2012; Hillocks, 2011; Newell, Beach, Smith, & VanDerHeide, 2011). Summarization requires the condensation of information to main ideas (A. L. Brown & Day, 1983; Westby, Culatta, Lawrence, & Hall-Kenyon, 2010). When the material to be summarized is presented in written text, summarization requires both reading comprehension and writing skill (Fitzgerald & Shanahan, 2000; Mateos, Martín, Villalón, & Luna, 2008). Both persuasive writing and summarization are featured in the Common Core State Standards for College and Career Readiness and in college student learning outcomes. In the current study, students were asked to respond to two writing prompts, one requiring a summary and the other a persuasive essay, based on written source text.

Self-efficacy for reading and writing. *Self-efficacy*, defined as the level of confidence a person has in his/her own ability to perform a challenging task (Bruning, Dempsey, Kauffman, McKim, & Zumbrunn, 2013), is an important construct in postsecondary education. This variable is associated with academic achievement and perseverance, including a tendency to increase effort or attempt new strategies in the face of academic difficulty. Conley and French (2014) added self-efficacy in an extension of Conley's (2007) basic model of college readiness.

Self-efficacy appears to mediate learning among college students (Kitsantas & Zimmerman, 2009), and statistically significant relationships between this construct and reading or writing achievement have been reported (MacArthur, Philippakos, & Graham, 2015; Martinez, Kock, & Cass, 2011; Pajares & Valiante, 2006; Proctor, Daley, Louick, Leidera, & Gardner, 2014). College developmental reading students report lower levels of self-efficacy than students not placed in developmental reading courses (Cantrell et al., 2013). Further, low-skilled postsecondary students may increase their self-efficacy for reading and writing tasks as their skills improve (Caverly, Nicholson, & Radcliffe, 2004; MacArthur & Philippakos, 2013; MacArthur, Philippakos, & Ianetta, 2015). For example, in an intervention study conducted with students in college basic writing courses, students who received explicit and structured instruction in writing strategies and self-regulation reported higher levels of writing self-efficacy than students who received traditional writing instruction (MacArthur, Philippakos, & Ianetta, 2015).

Teacher judgments of students' reading and writing skills. Students' self-efficacy reports may be subject to social desirability effects, or to effects of low metacognitive skills, which would create difficulties for self-reflection on skill levels. One of the aims of the current study is to assess the relationship between students' self-efficacy ratings on specific reading and writing tasks and the judgments of teachers on their students' ability to perform the same tasks. Discrepancies could result from inaccuracy on either students' or teachers' part, or both. Therefore, the results of this comparison cannot be interpreted with any certainty but could provide the basis for formulating hypotheses that could be tested in subsequent research. (A literature search indicated a lack of studies comparing student self-efficacy and teacher judgments on the same skills.)

Teachers have been found to be reliable judges of their students' reading ability (Ritchey, Silverman, Schatschneider, & Speece, 2015). An early review of studies of the relation between teacher ratings and student test scores reported correlations from $r = .28$ to $r = .86$, with a median correlation of $r = .62$ (Hoge & Coladarci, 1989). A more recent review corroborated this finding, with a mean effect size of .63 (Südkamp, Kaiser, & Möller, 2012). Teacher judgments of students' general writing ability were found to be moderate predictors of students' motivational beliefs about writing, which included self-efficacy (Troia, Harbaugh, Shankland, Wolbers, & Lawrence, 2013). However, it has also been reported that teacher ratings of reading skill are more reliable predictors of the performance of higher achieving students than of the performance of lower achieving students (Begeny, Krouse, Brown, & Mann, 2011; Feinberg & Shapiro, 2009). Further, despite statistically significant correlations between teacher judgments and students' reading scores, teachers tend to overestimate their students' actual ability, especially for students of average reading ability (Martin & Shapiro, 2011).

Students' insight via retrospective reports. First-person accounts of task performance may help in the interpretation of students' performance scores. For example, low scores may reflect low ability, lack of understanding of task requirements, lack of motivation, or some combination of these variables. For these reasons, student depictions of their own performance can provide important information on their level of college readiness. Two methods have been used to obtain first-hand accounts of performance: concurrent think-aloud statements and retrospective reports (Ericsson & Simon, 1993; Merchie & Van Keer, 2014). The former requires that the participant verbalize "online" the thoughts that are occurring and the strategies he or she is using while doing the task. This method is cognitively demanding, requires much training, and may be unreliable for participants who have low verbal or metacognitive skills and related difficulty in self-reflection. The latter, retrospective reports, involves interviews or questionnaires administered after a task is completed (Aghaie & Zhang, 2012; Harrison & Beres, 2007). Here, the participant is asked specific questions about how he/she interpreted the task and what he/she did to perform it. Although retrospective reports may be subject to problems related to inaccurate memory, difficulty reflecting on one's own process, and/or social desirability (with interviewee stating what he or she thinks the interviewer will value), the

advantages of obtaining first-person reports from students seemed to outweigh the disadvantages of this approach. Therefore, we employed retrospective reports in the current study to obtain students' insights into their performance on one of the assessment tasks.

Retrospective reports have been used to help explain performance on a wide range of reading and writing tasks. Most of the studies we identified in a literature search were conducted with children (e.g., Crammond, 1998; Farrington-Flint, Coyne, Stiller, & Heath, 2008; Farrington-Flint & Wood, 2007; Griva, Alevriadou, & Semoglou, 2012; Kwong & Varnhagen, 2005; Moore & MacArthur, 2012; Steffler, Varnhagen, Friesen, & Treiman, 1998), although several were conducted with college students, including both English language learners and typical university students (Chou, 2013; Kwong & Brachman, 2014; Plakans, 2008; Strømsø, Bråten, Britt, & Ferguson, 2013). No studies were identified where retrospective reports were obtained from postsecondary developmental education students.

4. Development of the Researcher-Practitioner Partnership

4.1 Characteristics and Benefits

The current study was conducted by a partnership between a university research center and two community colleges. Recent literature has identified researcher-practitioner collaboration as an important tool to support the development of effective policies and practices (Coburn, Penuel, & Geil, 2013; Torraco, 2014). Such collaboration may benefit both researchers and practitioners, given that participants not only take part in the investigation of issues and the application of strategies but also gain insight into each other's experiences and perspectives in their respective roles (Coburn et al., 2013). In addition, Torraco (2014) suggests that scholar-practitioner collaboration ultimately creates stronger programming as a result of collaborators bringing different sources of knowledge to the conversation. Specifically, this collaboration allows for the initial consideration of both research and practice as offering different aspects of knowledge that are valuable. The partnership takes into account different facets of knowledge, including not only knowledge that is produced through research but also information that is embedded in practice itself.

Coburn et al. (2013) identified five features that characterize researcher-practitioner partnerships in education and distinguished this form of collaboration from others. Researcher-practitioner partnerships are long-term. Both parties commit to working together for an extended period of time rather than through a single interaction. This feature allows both parties to address complex questions. Moreover, through a partnership, researchers and practitioners are able to develop familiarity and trust over time, which enable parties to resolve challenges that may come up. These partnerships focus specifically on a problem of practice, or on an issue that is relevant to the work of educational administrators or instructors and will lend itself to applied research; examples include instructional and curriculum design and student learning. Researcher-practitioner partnerships also involve interactions between both parties that consistently address issues both partners find important. These partnerships rely on strategies that are intentionally designed to facilitate mutually beneficial interactions, such as opportunities for codesigning studies. Finally, Coburn et al. (2013) suggest that the research produced from these partnerships is unique and extends beyond analysis that may be completed internally by educational institutions.

In education, researcher-practitioner partnerships support research and decision making related to various complex problems that occur in education settings (Torraco, 2014). Developmental education is one area in which these kinds of partnerships are underutilized. However, as noted by Torraco (2014), collaborations between educators and researchers in postsecondary education have been beneficial; not only has new knowledge been produced, but this knowledge has also been applied directly to practice. Torraco wrote specifically about the issue of remedial education and contended that given that remedial education is considered a complex problem of practice with many contentious issues, additional integration of researcher and practitioner perspectives would be particularly useful.

4.2 The Current Partnership

The current study aligns with both Coburn et al.'s (2013) partnership features and Torraco's (2014) vision for researcher-practitioner partnerships in remedial education. To investigate the literacy skills of developmental English students, the research partner, the Community College Research Center (CCRC) at Teachers College, Columbia University,

collaborated with practitioner partners, who were community college instructors and administrators in two colleges. The study was part of a long-term research alliance between CCRC and the community college system, situated in a southern state (the state and the participating colleges are anonymized). The system encompasses all of the community colleges in the state and has instituted centralized policies and prescribed student learning outcomes for developmental education. Within the last 10 years, the system has been implementing a statewide reform of developmental reading, writing, and mathematics courses. Importantly for the current study, the reading and writing courses have been combined in the form of single developmental English courses. Further, these integrated courses are taught in eight-week periods, replacing the prior 16-week courses.

CCRC has partnered with the community college system to explore the nature, implementation, and early outcomes of the reform. The current study centers on two community colleges in the system. Each of these colleges served both urban and suburban areas. Senior administrators at both colleges committed to participation in the partnership for a minimum of two years. This relatively long-term commitment allowed the researcher and practitioner partners to work together throughout the development of the research design and the data collection phases of the study.

The research conducted in this partnership focuses on student knowledge, competencies, and skills. Attention to these areas in developmental courses not only fills an important gap in the literature but also is useful within the context of a researcher-practitioner partnership. When CCRC researchers presented the concept underlying the present study to senior college administrators, they responded positively and agreed to participate because they recognized it as an opportunity to obtain findings that can inform curriculum refinement in their developmental reading and writing courses. More specifically, the practitioner partners indicated that improved understanding of the skills and knowledge students gain in developmental courses would inform decision making on curriculum and pedagogy in the future.

To leverage the partners' respective expertise and produce mutually beneficial research, the partnership developed a process of communication that brought the participating researchers and practitioners together. On-site interviews with instructors of both developmental and college-level courses as part of a larger project verified the

importance of written summarization and persuasive writing, the two tasks used in the current study. At the beginning of the learning assessment study, each college named a research liaison and two developmental English instructors to serve as lead faculty partners. These individuals provided on-site logistical support, such as recruiting students to participate in assessments. Further, early in the development of the partnership, the college liaisons articulated to CCRC their respective colleges' objectives for participating in the study. Both liaisons were deans, and they approached the study from the perspective of supporting the curricular and pedagogical development of their developmental English courses. Moreover, lead faculty allowed researchers to observe multiple class sessions in order to familiarize themselves with instructional practices, faculty styles, and student participation in classrooms in the study sites. Lead faculty and research liaisons also contributed to the selection and design of the assessment instruments used in the study.

There was a large amount of communication between the CCRC researchers and the college leads from the beginning of the development of the partnership. An important event in the collaborative process was a one-day retreat, which was held at one of the two colleges; the research team and the faculty leads and research liaisons from both colleges attended the retreat. The retreat included a broad discussion among participants of the study's goals and workshop sessions for in-depth discussion of the assessments. In particular, at the retreat, the research team worked closely with the practitioner partners to assess the appropriateness of the written text that the researchers proposed to use in the text-based writing tasks, the nature of the writing prompts, attributes of students' writing to be evaluated, and the assessment administration procedures. There was a great deal of discussion on the nature of the prompts, and especially on how critical thinking could be assessed. The wording of the prompts was carefully crafted in this collaborative procedure.

As the assessment was being developed by the partnership, two practical constraints became apparent. First, the instructors did not wish to give up classroom time for the research, given their need to meet curricular goals and prepare students for tests. Second, the instructors did not wish the students to be subjected to what they considered to be an excessive amount of testing. Keeping these constraints in mind, the partners agreed on an assessment battery consisting of a mixture of standardized and researcher-developed

measures and a student background questionnaire. Drafts of researcher-designed measures were submitted by the research team to the community college partners for review.

5. Method

5.1 Research Questions

The partnership's overall aim was to administer an assessment of the academic literacy skills of developmental reading and English students to determine how prepared they were for introductory college-level literacy demands. The study used mixed methods comprising quantitative measures of reading, writing self-efficacy, and teacher judgments and a qualitative analysis of retrospective reports. The central focus of the study was performance on text-based persuasive writing and summarization tasks. Both of these tasks required that the student read a printed text and then answer a question in writing. Interviews with faculty teaching introductory college-level disciplinary courses, conducted during visits to the two colleges that took place as the study was being planned, indicated that competence in such tasks was important in their classes. These tasks were designed expressly for the study and keyed to a specific statewide learning goal for developmental education students stating that students would be able to critically analyze texts at a level needed for college and careers. The tasks also reflected the fact that reading and writing were integrated in the developmental courses, and that there is a strong theoretical relation between the two skills (Fitzgerald & Shanahan, 2000).

The study sought to answer the following questions:

1. How close are students to being ready for introductory college-level reading and writing?

   Explanation: The study provides a snapshot of where students are toward the end of their eight-week integrated developmental reading and writing course, in terms of key skills they would need at the introductory college level. Students from both intermediate and top-level developmental courses were recruited for the assessment, and we expected students at the top level to be more ready for college literacy demands than students at the intermediate level. There are no clear external criteria in prior literature by which to determine definitively whether students are college-ready. Therefore, we interpreted our

assessment data based on 12th-grade end-of-year standardized test norms, as well as on the proficiency of research samples that were as similar as possible to the current participants.

2. What is the level of students' self-efficacy, and what are teachers' judgments, in relation to students' ability to perform introductory college-level reading and writing tasks?

    a. How confident are students and teachers in students' reading and writing ability?
    b. How close are students' self-efficacy ratings to their teachers' judgments of their reading and writing skills?

    Explanation: Self-efficacy, defined as the amount of confidence someone feels in his or her ability perform a demanding task, is an important construct in education. In this study, we are interested in knowing how confident students feel in their ability to read and write at the college level, and how close their levels of confidence are to their instructors' views of their skills.

3. What are the correlations between standardized reading and writing scores, self-efficacy ratings, teacher judgments, and text-based writing performance?

    Explanation: This question looks at how the various components we assessed "hang together" as a way of thinking about how one variable might predict another.

4. What are the contributions of standardized test scores, self-efficacy ratings, and teacher judgments to text-based writing measures, controlling for college attended?

    Explanation: Here, as with the correlations, we look at interrelationships between variables, but we include all variables in one equation in order to examine which ones are most important in explaining students' performance on text-based writing tasks (persuasive writing and summarization). The control variable was the college students attended, which was used in order to remove effects from attending one college or the other.

5. How do students conceptualize the demands of introductory college-level reading and writing tasks?

Explanation: We interviewed some of the participants in order to seek their perceptions of their ability on the written summarization task we used in the assessment. The interview was developed as a retrospective report in which students described their experiences of writing a summary shortly after completing the task. In particular, we were interested in whether the students understood the nature of summarization and the strategies they used to summarize a written text.

5.2 Participants and Setting

The participants were 211 students attending developmental education courses in two community colleges (which we refer to as College 1 and College 2) in a southern state. College 1 was situated in a mid-sized city and served an urban population. Enrollment in college-credit courses was 17,937; 53 percent of students were male, 52 percent were White, 23 percent were Black or African American, 53 percent were aged 24 or below, and 7 percent had registered for developmental reading and English courses. College 2 was located in a suburb of a small city and served an urban–suburban community. College-credit enrollment was 7,676, with 39 percent male students, 30 percent White students, 28 percent Black or African American students, and 12 percent registered for developmental reading and English courses.

According to centralized state policy, all developmental reading and writing instruction was integrated in single developmental reading and English courses, which were taught in an eight-week, compressed acceleration model (Edgecombe, 2011). This policy was part of an ongoing statewide restructuring of the state's developmental reading, writing, and mathematics program. At the time of data collection, College 1 was integrating reading and writing and using the accelerated time frame for the first time, and College 2 had been doing this for several years. There were three levels of the integrated developmental reading and English course, and study participants attended the intermediate and top levels. At both colleges, the largest proportion of developmental reading and English enrollments (65 percent at College 1 and 56 percent at College 2) were in the top-level courses, which are one level below college-credit English. The state mandated that the student learning goals listed in Box 1 be addressed at all levels of the developmental reading and English curriculum.

Box 1
State Learning Goals for Integrated Developmental Reading and Writing

- Demonstrate the use of reading and writing processes.
- Apply critical thinking strategies in reading and writing.
- Recognize and compose well-developed, coherent, and unified texts.

Within the sample of 211 students, 123 attended College 1, and 88 attended College 2. At the time of data collection, the students were nearing the end of their course, and those passing the top-level course would be considered ready for college reading and writing. Instructors reported informally to the researchers that almost all of the study participants were expected to pass, as students who were failing their assignments had previously withdrawn from the course.

In the whole sample, 54 percent of the students were Black or African American, 64 percent were female, and 75 percent spoke English as a native language. Students with other native languages were all fluent English speakers. The mean age was 24.55 years (SD = 10.66), and 71 percent of the students were aged 18 to 24 years. Further information on student background for the whole sample and by college is shown in Table 1. For the sample as a whole, mean reading and writing scores translate to the 22nd percentile for the Comprehension subtest of the Nelson-Denny Reading Test and the 27th percentile for the Woodcock-Johnson III (WJ III) Writing Fluency subtest, using as a reference group 12th graders at the end of the school year.

Table 1
Student Background Characteristics

Variable	College 1 (*n* = 123)		College 2 (*n* = 88)		Total Sample (*N* = 211)	
	n	%	*n*	%	*N*	%
Demographics						
Age in years: 18	35	28	25	28	64	30
Age in years: 19–24	25	20	14	16	87	41
Age in years: 25+	32	26	29	33	61	29
Female	76	62	60	68	136	64
Black/African American	44	36	71	81	115	54
White	44	36	12	14	56	26
Primary language growing up: English	77	62	80	91	158	75
Responsible for children at home	36	29	30	34	66	31
Employed	79	64	45	51	123	59
Education						
High school diploma	105	85	70	80	175	83
GED	14	11	13	15	27	13
College-level diploma or certificate	6	5	11	13	17	8
Previous college reading or writing course	50	40	38	43	88	42
Academic behaviors						
Never used tutoring	49	40	45	51	94	44
Used tutoring once or twice	46	37	21	24	67	32
Used tutoring three times or more	19	15	10	11	29	14
Never asked for extra help	35	28	28	32	63	30
Asked for extra help once or twice	53	42	32	36	85	40
Asked for extra help three times or more	26	21	16	18	42	20
Motivation						
Gave best effort on research tasks (agree or strongly agree)	96	77	78	88	159	75
Highly motivated for research tasks (agree or strongly agree)	92	74	67	77	174	82
Educational goals						
Transfer to four-year college	69	56	41	47	110	52
Earn associate degree	68	55	41	47	109	51
Earn diploma or certificate	13	11	17	Z	30	14
Take some classes	9	7	4	5	13	6

5.3 Assessment Tasks

The assessment consisted of seven measures, listed in Box 2. Two were standardized tests, which were administered to determine students' level of general reading and writing skill and which served as covariates in the data analysis. The other measures, which were designed for the study, were a text-based summarization task, a text-based persuasive writing task, a self-efficacy questionnaire, a teacher judgment questionnaire, and a retrospective report interview.

In addition, a project-designed student background questionnaire asked about demographics; current employment; academic background; and, using a 5-point Likert-type scale, students' motivation and effort on the assessment. These questions were asked to try to ascertain whether students' performance on the measures reflected a true intent to produce their best work; as reported in Table 1, the large majority of responses suggest that it did.

Box 2
Instrumentation

Standardized measures

- Reading comprehension: Nelson-Denny Reading Test, Comprehension subtest (J. I. Brown, Fishco, & Hanna, 1993, Form H)
- Sentence writing: Woodcock-Johnson III Tests of Achievement, Writing Fluency subtest (Woodcock, McGrew, & Mather, 2001)

Project-designed measures

- Text-based writing
    - Persuasive writing: Read a newspaper article and answer a prompt on a controversy emerging in the article
    - Written summarization: Read and write summary of different newspaper article
- Self-efficacy for text-based writing
- Teacher judgments keyed to the student self-efficacy items and collected online using Qualtrics software
- Student retrospective reports on the text-based written summarization task

Reading comprehension. The Comprehension subtest of the Nelson-Denny Reading Test, Form H (J. I. Brown et al., 1993), was administered to measure students' ability to understand printed text. This is a standardized measure in which the test taker is given 20 minutes to respond to 38 multiple-choice factual and inferential questions based on seven reading passages on assorted unrelated topics. Raw scores were used in the analysis, doubled following instructions in the test manual. The test is normed for grades 9 through 16. The test's publisher reports Kuder–Richardson Formula 20 reliability coefficients of .85 to .91 for the Comprehension subtest but does not provide information on validity. However, the measure has reasonable face validity for screening general reading skills (Corkill, 2007).

Sentence writing. The Writing Fluency subtest of the WJ III Tests of Achievement (Woodcock et al., 2001) is a standardized measure of general writing skill. Specifically, it tests students' ability to formulate and write sentences quickly. The test presents 40 items, each consisting of three words. On each item, all three words must be included in a grammatically correct sentence. Students are given 7 minutes to complete the test. Raw scores were used in the analysis. The WJ III battery is normed for ages 2 through 90. The median score reliability, using a Rasch procedure appropriate for speeded measures, is .88 (reliabilities of .80 and above interpreted as desirable; Schrank, McGrew, & Woodcock, 2001). Although the Writing Fluency subtest is hardly an authentic test of writing ability, it has the advantages of being scored according to objective criteria, requiring short testing time, and having been used in previous research on developmental education students (MacArthur, Philippakos, & Ianetta, 2015), where a statistically significant relationship was found between the measure and the quality of authentic essays.

Text-based writing. Text-based writing ability was assessed with two 30-minute tasks using two articles from the newspaper *USA Today*. The articles were selected to correspond to topics taught in high-enrollment, introductory-level, college-credit general education courses with significant reading and writing requirements in the two colleges. An inspection of enrollments using institutional data from the two sites indicated that the highest enrollments in courses meeting these criteria were in psychology and sociology. The liaisons at the two sites indicated that the participants had not yet taken the courses. Prior interviews with content-area faculty had indicated that newspaper articles were used

regularly to supplement the use of textbooks in these courses. The tables of contents of the introductory-level psychology and sociology textbooks used at the two colleges were used as the basis of a search for appropriate newspaper articles. The criteria for the selection of articles were relevance to topics listed in the tables of contents, word count, and a level of readability that was feasible for the participants. Introductory college textbooks tend to be written at the 12th-grade readability level, although developmental education texts are generally written at a lower level (Armstrong et al., 2015).

The two articles selected were on the psychology topic of stress experienced by teenagers and the sociology topic of intergenerational tensions in the workplace. The psychology topic was used for the persuasive essay, and the sociology topic was used for the written summarization task. Text characteristics are shown in Table 2.

Table 2
Text-Based Writing Task Text Characteristics

Subject Area	Topic	Flesch–Kincaid Grade Level	Lexile	Word Count
Psychology	Teen stress	10.5	1250L	650
Sociology	Intergenerational conflict	11.1	1340L	676

As indicated in the section above on the development of the researcher-practitioner partnership, the practitioner partners wished to limit the amount of time students would spend on the assessment. An inspection of textbook chapters indicated that, in order to maintain coherence and meaning of the text to be read, several pages would have to be presented, which would require more time than was considered feasible by the practitioner partners. As an alternative, it was decided to use newspaper articles. Further, to meet the time requirements considered feasible for the study, it was necessary to use relatively short articles. Because newspaper articles that were considered readable by the target population and that also corresponded to introductory psychology and sociology courses were too long for the time to be given for the tasks, research staff reduced their word length by eliminating several paragraphs in each article. The deleted paragraphs presented examples to illustrate main points in the articles and did not add new meaning. Two teachers with experience in reading instruction read the reduced-length articles and verified that neither cohesion nor basic meaning had been lost as a

result of the reduction. Word counts for the articles were 650 for the psychology text and 676 for the sociology text.

The readability of the two texts was measured using Flesch–Kincaid grade levels, found in the Microsoft Word software program, and Lexile measures. As shown in Table 2, the texts had 10th–11th grade Flesch–Kincaid reading levels and were at 1250–1340 Lexile levels, interpreted as corresponding to an approximately 12th-grade (end-of-year) level. In the context of Common Core State Standards, high school students who can comprehend text at a 1300 Lexile level are considered ready for college and career-related reading tasks (National Governors' Association & Council of Chief State School Officers, 2010, Appendix A).

Text-based persuasive writing. Students were asked to read a newspaper article and write an essay expressing their opinion (i.e., write a persuasive essay) on a controversy discussed in the article. The prompt, which was developed in collaboration with the practitioner partners, asked the students to read the article and express their opinion on the controversy in their own words, using no quotations. The instructions directed the students to pretend that they were trying to persuade a friend to agree with them. The directions also stated that students could mark the article in any way and that they could use a dictionary during the task.

Text-based summarization. The text-based summarization task required students to read a second newspaper article and summarize it in one or two paragraphs using their own words, again with no quotations. On the recommendation of the practitioner partners, the prompt also asked students to state how the information in the article could be applied in a job setting and to support their answer with examples from the reading. This part of the prompt was intended to capture critical thinking, a state learning goal for developmental education. As with the persuasive writing task, students were permitted to use a dictionary and mark the source text as they wished.

Student self-efficacy. On a scale we called the Student Self-Reflection Questionnaire, based on the work of MacArthur, Philippakos, and Ianetta (2015), students were asked to respond to 16 questions that asked them to rate the level of confidence they felt in their ability on the text-based tasks and related skills. The ratings were obtained prior to the administration of the summarization and persuasive writing

tasks so that they could be used predictively in the analysis. The instructions informed respondents that they would shortly be asked to read two newspaper articles and write a summary and persuasive essay based on them. They were directed to rate their confidence on these tasks by selecting, for each of the 16 questions, one point on a 100-point scale reflecting their level of confidence. Participants had to circle one of 11 points on the scale: (0, 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100). Examples of points on the scale were provided in the instructions: 0 ("you are sure you cannot do it"), 50 ("there is an equal chance that you can do it or not do it"), and 100 ("you are sure you can do it"). The same 16 questions were used for teacher judgments, as described below. However, the wording of the student items was simplified in order to ensure comprehension.

Teacher judgments. A teacher judgment scale was developed for the study based on work suggesting that teacher judgments are predictive of student performance (Hoge & Coladarci, 1989; Speece et al., 2010; Südkamp et al., 2012; Troia et al., 2013). Using the online Qualtrics platform, instructors were asked to provide ratings on the same items for which students had rated their self-efficacy, using the same 11 points on a 100-point scale. Thus, each teacher judgment was keyed to a self-efficacy item. For instance, for a student self-efficacy item that stated, "I can read the articles carefully and form my own opinion about the issues discussed," the corresponding teacher judgment item stated, "The student can read the passages carefully and think critically about the ideas discussed." The full list of student and teacher questions is shown in Appendix B.

Retrospective reports. Retrospective reports were obtained from some of the participants in order to sample their understanding of task instructions and obtain their accounts of strategies used to complete the task. The retrospective report focused solely on the summarization task. In a one-hour interview, participants were asked 24 questions and related probes concerning their performance on that task. The questions fell into several categories: (1) what the student thought and did during the task, (2) planning what to write, (3) writing, (4) comprehension of the source text, and (5) knowledge and experience of the requirements of text-based summarization.

5.4 Assessment Procedure

The data were collected at the two colleges in November 2014 after researchers piloted the assessment with another group of students at the colleges in April of the same

year. Participants were recruited using flyers and in-class announcements by teachers, facilitated by college liaisons. After a pool of students had been recruited, participants were selected based on confirmation that they were attending the intermediate or top-level integrated developmental reading and English course and were available for the 2.5 hours required for the assessment.

The assessment was conducted during non-class hours and coordinated on-site by the college liaisons. Participants provided signed consent using a form approved by the Teachers College Institutional Review Board. In the course of obtaining consent, research staff emphasized the confidentiality of all data collected, that participation and performance on the assessment had no bearing on students' standing in their classes, and that they could withdraw from the study at any time without penalty.

The assessment tasks were administered in a 2.5 hour session with rest breaks, in the following fixed order: Student Self-Reflection Questionnaire (self-efficacy), text-based writing task 1 (persuasive essay), WJ III Writing Fluency subtest, text-based writing task 2 (summarization), Nelson-Denny Comprehension subtest, and student background questionnaire. The tasks were administered by CCRC research staff in classrooms at the two colleges. A member of the research team led each session using a script containing task instructions. All tasks except the retrospective report were administered to groups. All tasks, including the text-based writing tasks, were completed using pen and paper.

The retrospective reports were obtained individually from 28 of the 211 participants in one-hour interviews after completion of the group assessment. The interviewees were selected based on their availability and willingness to devote an additional hour to the research. The interviews were conducted by CCRC staff in offices or classrooms at the colleges and were audio-recorded and later transcribed. Incentives in the form of gift cards were given to participants upon completion of each phase of the assessment (group testing and individual interview as applicable).

To obtain the teacher judgments, the participants' developmental course instructors were contacted by the college liaisons prior to the assessment to inform them of the purpose of the study and to request ratings on their students' class performance. Upon completion of the group assessment, the research team sent emails to the

instructors containing instructions and a Qualtrics web link for submission of their judgments. Attached to the emails were copies of the reading and writing tasks on which the teacher judgments focused. The instructors received monetary compensation for submitting their judgments.

5.5 Scoring

The persuasive essays and summaries were word processed, correcting for spelling, capitalization, and punctuation, in order to reduce bias in scoring (Graham, 1999; MacArthur & Philippakos, 2010; Olinghouse, 2008). Grammatical errors were not corrected. Four scores were obtained for the persuasive essays. The first score came from a 7-point holistic persuasive quality rubric (based on MacArthur, Philippakos, & Ianetta, 2015). While scoring, raters were asked to bear in mind the clarity of expression of ideas, the organization of the essay, the choice of words, the flow of language and variety of sentences written, and the use of grammar. Examples of two score points on the 7-point holistic persuasive quality scale are shown in Box 3.

Box 3
Examples of Holistic Persuasive Quality Score Points

Score = 3
Essay has topic and a few ideas but little elaboration. Ideas not explained well or somewhat difficult to understand. Source text not mentioned or referred to vaguely. Less important details rather than main ideas from the source text used to support argument. Some ideas conveyed inaccurately. If personal experience mentioned, largely irrelevant or mostly unclear. Organization may be weak. Essay may lack introduction and transitions among ideas. Word choice may be repetitive or vague. Sentences may be simple or lack variety. Errors in grammar and usage.

Score = 5
Clear topic with related ideas supported with details and some elaboration. Source text mentioned explicitly. Some main ideas from text used to support the argument. Most of ideas from source text conveyed accurately. If personal experience mentioned, mostly relevant and clear. Essay well organized, with introduction, sequence of ideas with some transitions, and conclusion. Word choice generally appropriate and some variety of sentences. Occasional minor errors in grammar or usage.

The second score was a count of persuasive essay parts included in the essay (adapted from Ferretti, MacArthur, & Dowdy, 2000). Based on prior research (Gil, Bråten, Vidal-Abarca, & Strømsø, 2010), each essay was first parsed into idea units, defined as follows:

> An idea unit contained a main verb that expressed an event, activity, or state. If an utterance had two verbs and one agent, it was treated as having two separate idea units. Infinitives and complements were included with the main verb. (Magliano, Trabasso, & Graesser, 1999, p. 44)

Each parsed unit was then labeled with one of the following codes, taken from Ferretti et al. (2000):

- **P:** proposition, or statement of belief or opinion;
- **R:** reason for the position stated;
- **E:** elaboration of proposition or reason;
- **AP:** alternative proposition, or counterargument;
- **AR:** reason for alternative proposition;
- **RB:** rebuttal of the counterargument;
- **C:** concluding statement; or
- **NF:** nonfunctional units, defined as repetitions or information not relevant to the prompt.

Almost all of the functional units were propositions, reasons, elaborations of propositions or reasons, or conclusions, with very few counterarguments or rebuttals.

The third score was the number of academic words contained in the writing sample (Lesaux, Kieffer, Kelley, & Harris, 2014; Olinghouse & Wilson, 2013). This is a measure of vocabulary usage, operationalized as the number of words in a writing sample that appear frequently in academic texts but are not specific to any specific subject area (Lesaux et al., 2014). Examples of academic words are *circumstances*, *category*, *debate*, *demonstrate*, *estimate*, *interpret*, and *guarantee* (Coxhead, 2000). The number of academic words contained in the text-based summaries and persuasive essays, expressed as a percentage of the number of words written, was obtained from the automated vocabulary profiler software program "VocabProfile" (http://www.lextutor.ca/vp/eng/). This profiler computes the

percentage of words that occur on the Academic Word List constructed by Coxhead (2000), which contains groups of words—specifically, 570 word families—covering approximately 10 percent of the words found in academic materials that are *not* among the 2,000 most frequent words in the English language. This vocabulary measure reflects the assumption that less mature writing contains a predominance of highly frequent words and fewer low-frequency words (McNamara, Crossley, & McCarthy, 2010; Uccelli, Dobbs, & Scott, 2013).

The fourth score was the length of the essay, or the number of words written. This measure is frequently used in writing research, and low-achieving students often produce very short compositions containing a small amount of information (Doolan, 2014; Nelson & Van Meter, 2007; Puranik, Lombardino, & Altmann, 2008). MacArthur, Philippakos, and Ianetta (2015) found that teaching developmental education students to use a writing strategy resulted in significantly longer compositions.

The written summaries were also scored on four components. First was a count of main ideas from the source text that the student included in the writing sample (Perin, Bork, Peverly, & Mason, 2013; Perin, Keselman, & Monopoli, 2003). Since a summary should contain the gist, or main ideas, of the source, a count of the main ideas in the newspaper article that are included in a student's written summary is one way of assessing its quality. To identify the main ideas in the newspaper article, the first author and two research assistants who were experienced English language arts teachers each independently read the article and listed the main ideas. The three individuals then discussed each idea and came to consensus on which ideas to retain, eliminate, or add to the list. For the analysis, the number of main ideas found in each summary was expressed as a proportion of the nine main ideas in the source text.

The second score represented the quality of summary, using an analytic summarization quality rubric with four 4-point ratings for a maximum score of 16 (based on Westby et al., 2010, p. 283). Ratings were made for four aspects of summary quality: (1) topic/key sentence, main idea; (b) text structure; (c) gist, content (quantity, accuracy, and relevance); and (d) sentence structure, grammar, and vocabulary. Score points for the elements of *topic/key sentence, main idea* and *gist, content (quantity, accuracy, and relevance)* are shown in Box 4 as an example. Box 5 describes a score of 3 on each of the four elements.

Box 4
Examples of Summary Quality Score Points (adapted from Westby et al., 2010)

Topic/key sentence, main idea

0 (None): Statements do not link to central topic.

1 (Little/few/some): Ideas link to central topic, but no topic/key sentence brings ideas together.

2 (Many/most): Topic/key sentence states some aspect of the content, but statement is vague.

3 (All): Topic/key sentence is minimal but clear.

4 (Best): Topic/key sentence is comprehensive.

Gist, content (quantity, accuracy, and relevance)

0 (None): Statements are not related to the passage or do not communicate information from the passage.

1 (Little/few/some): Some information from the passage is included, but some important ideas are missing; some ideas may be irrelevant or inaccurate.

2 (Many/most): Most information from the passage is included; some ideas may be irrelevant or inaccurate; some information/ideas are missing.

3 (All): All relevant information from the passage is included.

4 (Best): All relevant ideas from the passage are included, accurately represented, and appropriately elaborated.

Box 5
Example of One Analytic Score Point on Each Element

Score of 3 on each of four quality indicators

Topic/key sentence, *main idea*: Topic/key sentence is minimal but clear.

Text structure: Logical flow of information.

Gist, content (quantity, accuracy, and relevance): All relevant information from passage included.

Sentence structure, grammar, and vocabulary: All sentences complete, some elaboration and/or dependent clauses, generally appropriate word choice, some variety of sentence structure, occasional minor errors of grammar and/or usage.

The third and fourth scores on the summarization task were the number of academic words used in the summary and word count, as obtained for the persuasive essays above.

All of the measures were scored by three trained research assistants who were experienced in literacy assessment but, to avoid bias in scoring, were unfamiliar with the goals of the project. The Nelson-Denny Comprehension and WJ III Writing Fluency subtests were scored following instructions in the test manuals. The WJ III Writing Fluency subtest and text-based writing protocols were scored by one of the research assistants, and a second assistant scored one third (70) of the writing samples for interrater reliability.

Each of the three assistants was responsible for scoring one of the full sets of writing samples. The assistants were trained by the first author. The rubric and anchor papers obtained from the pilot study were discussed, and the assistants practiced scoring sets of 10 protocols until they reached a criterion of 80 percent exact agreement. The percentage of interscorer agreement and Pearson product-moment correlations are shown in Table 3. Where there was not exact agreement, the score entered for analysis was the average of the two scores.

Table 3. Interscorer Agreement and Interrater Reliability

	% Interscorer Agreement			
Measure	**Exact**	**Within 1 Point**	**Within 2 Points**	**Interrater Reliability (*r*)**
WJ III Writing Fluency subtest raw score	40	79	87	.98
Essay: Number of functional elements	43	68	79	.92
Essay: Persuasive quality score	77	100		.85
Summary: Number of main ideas	14	63		.72
Summary: Quality score	23	44	71	.85

5.6 Coding of Retrospective Reports

The audio-recordings of the retrospective reports were transcribed and checked for accuracy. A codebook was drafted and revised by the first author and a trained doctoral assistant during the pilot study. To create the codebook, an initial set of codes was generated based on identification of themes in four interview transcripts obtained in the pilot. These codes were applied and then revised, and in some cases combined and eliminated, resulting in a total of 27 codes. The ATLAS.ti qualitative software program (http://atlasti.com/) was used for coding. Examples of codes are shown in Box 6. The 28

transcripts in the current full study were coded by two assistants, different from those who scored the writing samples, trained by the first author. After training, 15 of the transcripts were coded by both assistants. The first author discussed differences in coding with the assistants to identify differences in interpretation of interviewees' statements and clarify the meaning of the codes. Most differences were resolved through discussion.

Box 6
Retrospective Reports: Examples of Codes

- Appropriate reading strategy for the task
- Inappropriate reading strategy for the task
- Correct definition of summary
- Correct interpretation of information in the source text
- Correct understanding of the task
- Personal reaction to the source text
- Positive self-evaluation of task performance

6. Results

6.1 General Reading and Writing Ability

General reading and writing ability, as measured by the Nelson-Denny Reading Test Comprehension subtest (J. I. Brown et al., 1993) and the WJ III Writing Fluency subtest (Woodcock et al., 2001), is summarized for the whole sample and by college and developmental level in Tables 4 and 5. Raw scores were used in the analysis. Data on developmental course level were obtained from the student background questionnaire. The sample is slightly smaller in Table 5 because some participants did not provide information on their developmental level. For the whole sample, the mean raw scores on the Nelson-Denny Comprehension and WJ III Writing Fluency subtests corresponded to the 22nd and 27th percentile ranks, respectively, for 12th graders at the end of the school year. Across the whole sample ($N = 211$), there was a moderate but statistically significant correlation between these two tests ($r = .34$, $p < .01$).

Table 4
General Reading and Writing Scores: Whole Sample and by College (*N* = 211)

Test	College 1 *M* (*SD*)	College 2 *M* (*SD*)	Total Sample *M* (*SD*)
Nelson-Denny Comprehension subtest raw score (max = 76)	31.08 (13.16)	26.84 (10.52)	29.32 (12.29)
Nelson-Denny Comprehension subtest scale score (test mean = 200)	180.97 (19.65)	173.83 (16.49)	177.98 (18.69)
WJ III Writing Fluency subtest raw score (max = 40)	22.40 (5.73)	22.64 (4.49)	22.50 (5.25)
WJ III Writing Fluency subtest standard score (test mean = 100)	86.29 (16.98)	86.99 (13.33)	86.58 (15.55)
n	123	88	211

Table 5
General Reading and Writing Scores by College and Developmental Level (*n* = 194)

Test	College	Course Level	*n*	*M*	*SD*	*t*	*df*	*p*
Nelson-Denny Comprehension subtest (max = 76)	1	Intermediate	20	30.70	14.45	.109	116	.914
		Top	98	31.04	12.43			
	2	Intermediate	29	23.79	9.22	1.65	74	.103
		Top	47	27.83	11.00			
WJ III Writing Fluency subtest (max = 40)	1	Intermediate	20	23.30	5.78	.683	116	.496
		Top	98	22.34	5.74			
	2	Intermediate	29	22.86	4.83	.462	73	.645
		Top	46	22.35	4.61			

There was a statistically significant difference between the colleges on the Nelson-Denny Comprehension scores, with higher scores at College 1 ($t = 2.5$, $df = 210$, $p < .01$), but no difference in WJ III Writing Fluency scores between the college ($t = .33$, $df = 209$, $p = .744$). There were no statistically significant differences between the performance of students in the two developmental course levels on either test at either college (means and *t*-tests shown in Table 5). Thus, an unexpected finding of this study is that general reading and writing skills, as measured by the standardized tests, did not show reliable differences as a function of developmental level; in other words, the students in the intermediate and top levels were reading and writing at approximately the

same level. However, there was a trend (not statistically significant) for the reading scores to be higher among students in the top level course, but only at College 2.

6.2 Text-Based Writing

As described in the method section, text-based writing was assessed on eight variables (four for the persuasive essay and four for the written summary). Scores on these measures are summarized for the whole sample and by college and developmental course level in Tables 6–9. The sample sizes in the tables showing results by developmental level do not sum to the size of the whole sample because, as mentioned above, some students did not provide course information. Examples of students' written work can be found in Appendix A.

Comparison of colleges. There were statistically significant differences between the colleges on three of the eight text-based writing variables (see Tables 6 and 7): the quality of the summary ($t = 2.830$, $df = 210$, $p = .005$) and the percentage of academic words used in the summary ($t = 3.058$, $df = 210$, $p = .003$) and in the essay ($t = 5.567$, $df = 210$, $p = .000$).

Table 6
Text-Based Persuasive Writing Performance by College and for Total Sample

	College 1 (n = 123)		College 2 (n = 88)		Total Sample (N = 211)	
Variable	***M***	***SD***	***M***	***SD***	***M***	***SD***
Proportion functional elements of total elements	.52	.27	.54	.24	.53	.26
Persuasive quality score (holistic score, max = 7)	2.63	.80	2.51	.80	2.58	.80
Word count	188.41	77.64	177.75	67.10	183.99	73.47
% academic words	5.00	3.10	2.89	2.07	4.13	2.91

Table 7
Text-Based Written Summarization Performance by College and for Total Sample

	College 1 (*n* =123)		College 2 (*n* = 88)		Total Sample (*N* = 211)	
Variable	***M***	***SD***	***M***	***SD***	***M***	***SD***
Proportion of main ideas from source text	.26	.19	.22	.19	.24	.19
Summary quality score (analytic score, max = 16)	8.52	2.77	7.47	2.45	8.09	2.64
Word count	97.41	42.02	94.86	40.99	96.36	41.56
% academic words	5.00	3.28	3.71	2.588	4.46	3.07

Comparison of developmental course levels. There were no statistically significant differences between course levels on any variable in College 1 (see Table 8). On the essay, scores for the top level of the course were actually lower than for the intermediate level, although the differences were not statistically significant. On the summary, scores were higher for the top level on two of the variables but, again, the differences were not significant.

At College 2, there were statistically significant differences by course level on several variables in the expected direction (with lower performance at the intermediate level than at the top level). As shown in Table 9, significant differences were seen on four of the eight variables: essay quality ($t = 3.48$, $df = 74$, $p = .001$), essay word count ($t = 3.01$, $df = 74$, $p = .004$), summary quality ($t = 2.92$, $df = 74$, $p = .005$), and summary word count ($t = 2.25$, $df = 74$, $p = .027$). Further, the difference for the percentage of academic words in the essay approached conventional levels of statistical significance ($t = 1.87$, $df = 74$, $p = .065$). Overall, there is a clear difference between intermediate and top-level students at College 2 but not at College 1.

Table 8
Text-Based Writing Performance by Developmental Course Level: College 1 (*n* = 118)

Variable	Course Level	*n*	Mean	*SD*	*t* (*df* = 116)	*p*
Essay						
Proportion functional elements of total elements	Intermediate	20	.57	.30	.729	.468
	Top	98	.52	.27		
Persuasive quality score (holistic score, max = 7)	Intermediate	20	2.80	.70	.999	.320
	Top	98	2.61	.80		
Word count	Intermediate	20	193.05	56.84	.394	.694
	Top	98	186.20	73.21		
% academic words	Intermediate	20	5.53	3.96	.801	.425
	Top	98	4.91	2.97		
Summary						
Proportion of main ideas from source text	Intermediate	20	.19	.12	1.720	.088
	Top	98	.27	.20		
Quality score (analytic score, max = 16)	Intermediate	20	7.80	1.91	1.319	.190
	Top	98	8.67	2.83		
Word count	Intermediate	20	95.90	51.51	.190	.850
	Top	98	97.84	39.21		
% academic words	Intermediate	20	5.05	3.41	.073	.942
	Top	98	4.99	3.30		

Table 9
Text-Based Writing Performance by Developmental Course Level: College 2 (*n* = 76)

Variable	Course Level	*n*	Mean	*SD*	*t* (*df* = 74)	*p*
Essay						
Proportion functional elements of total elements	Intermediate	29	.56	.23	.935	.353
	Top	47	.51	.23		
Persuasive quality score (holistic score, max = 7)	Intermediate	29	2.10	.67	3.481	.001
	Top	47	2.72	.80		
Word count	Intermediate	29	151.34	51.86	3.011	.004
	Top	47	198.38	73.52		
% academic words	Intermediate	29	2.32	1.89	1.872	.065
	Top	47	3.23	2.15		
Summary						
Proportion of main ideas from source text	Intermediate	29	.18	.18	1.637	.106
	Top	47	.26	.19		
Quality score (analytic score, max = 16)	Intermediate	29	6.45	2.41	2.921	.005
	Top	47	8.11	2.39		
Word count	Intermediate	29	83.97	35.80	2.252	.027
	Top	47	105.48	42.84		
% academic words	Intermediate	29	3.41	2.19	1.013	.314
	Top	47	4.04	2.88		

6.3 Self-Efficacy Ratings and Teacher Judgments

The self-efficacy ratings and teacher judgments were each made using 11 points on a 100-point scale (0, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100), indicating the degree of confidence in the student's ability to perform various academic reading and writing tasks. For each student, a mean self-efficacy rating was computed over the 16 items on the Student Self-Reflection Questionnaire, and a mean teacher judgment was computed over the 16 items on the teacher judgment scale. For the sample of students who had complete data (all items answered on both the self-efficacy and teacher judgment measures, $n = 162$), there was a moderate, statistically significant correlation between the mean self-efficacy ratings and the mean teacher judgments ($r = .30$, $p < .01$). Mean self-efficacy ratings and teacher judgments for each item for the full sample, whether or not

data were complete ($N = 211$), are shown in Appendix B. The mean self-efficacy ratings and teacher judgments by college and developmental course level and for the whole sample are shown in Tables 10–12.

Comparison of colleges. For the students with complete data ($n = 162$), the colleges did not differ on self-efficacy ratings ($t = .915$, $df = 160$, $p = .362$) or teacher judgments ($t = .560$, $df = 160$, $p = .577$).

Table 10
Self-Efficacy Ratings and Teacher Judgments: Whole Sample and by College

Variable	College 1 *M* (*SD*)	College 2 *M* (*SD*)	Total Sample *M* (*SD*)
Mean self-efficacy rating	80.92 (12.51)	82.67 (10.74)	81.60 (11.85)
Mean teacher judgment	71.34 (15.35)	72.62 (12.16)	71.84 (14.17)
n	99	63	162

Comparison of developmental course levels. In comparing the developmental course levels, we report findings for students for whom we have complete self-efficacy and teacher judgment data, and for whom we have course numbers ($n = 152$). As shown in Tables 11 and 12, there were no statistically significant differences on either variable between developmental course levels at either college. At College 1, students in the top-level course had higher self-efficacy, but the difference was not statistically significant. At College 2, self-efficacy was slightly higher for the intermediate students than for the top-level students (a counterintuitive trend), and teacher judgments were higher for the students in the top level than for intermediate students, but again, these differences were not statistically significant.

Table 11
Self-Efficacy Ratings and Teacher Judgments by Developmental Course Level: College 1 (*n* = 96)

Variable	Course Level	*n*	Mean	*SD*	*t* (*df* = 94)	*p*
Mean self-efficacy rating	Intermediate	16	76.78	14.94	1.518	.132
	Top	80	81.87	11.67		
Mean teacher judgment	Intermediate	16	72.66	19.56	.3256	.746
	Top	80	71.32	14.01		

Table 12
Self-Efficacy Ratings and Teacher Judgments by Developmental Course Level: College 2 (*n* = 56)

Variable	Course Level	*n*	Mean	*SD*	*t* (*df* = 54)	*p*
Mean self-efficacy rating	Intermediate	22	84.26	9.37	1.094	.279
	Top	34	81.36	9.92		
Mean teacher judgment	Intermediate	22	69.72	11.16	1.581	.120
	Top	34	74.85	12.31		

6.4 Predicting Text-Based Writing Ability From Skills, Self-Efficacy Ratings, and Teacher Judgments

In this section, we examine the relationships between the eight measures of text-based writing ability and the standardized measures of reading and writing skill (Nelson-Denny Reading Comprehension and WJ III Writing Fluency subtests), self-efficacy ratings, and teacher judgments. These relationships are investigated using correlations and hierarchical regressions. We included college but not developmental course level as a covariate in the hierarchical regressions, since, as reported above, the latter did not show statistically significant differences on either of the standardized tests. Further, we included in the regressions only students with complete self-efficacy and teacher judgment data ($n = 162$).

Relationship between standardized test scores and text-based writing measures. Intercorrelations between the standardized test scores and text-based writing measures are shown in Table 13. Statistically significant correlations are listed in Box 7. Although there were numerous significant correlations, they were moderate at best. Notable findings include the correlations between the standardized reading test and the proportion of main ideas from the source text included in the summary ($r = .29$, $p < .01$) and between the word count of the writing sample and the essay and summary quality scores ($r = .51$, $p < .01$ for essay quality and $r = .50$, $p < .01$ for summary quality).

Table 13
Intercorrelations Between Standardized Test Scores and Text-Based Writing Scores

Variable	1	2	3	4	5	6	7	8	9	10
1. Nelson-Denny Comprehension	1									
2. WJ III Writing Fluency	.34**	1								
3. Essay: Proportion functional elements	.14*	.23**	1							
4. Essay: Persuasive quality score	.16*	.28**	.14*	1						
5. Essay: Word count	.17*	.20**	-.09	.51**	1					
6. Essay: % academic words	.13	.03	.01	.57**	.02	1				
7. Summary: Proportion of main ideas	.29**	.10	.10	.10	.15*	.12	1			
8. Summary: Quality score	.31**	.14*	.10	.16*	.23**	.13	.61**	1		
9. Summary: Word count	.09	-.01	-.06	.19**	.32**	.09	.47**	.50**	1	
10. Summary: % academic words	.15*	-.01	.11	.04	-.01	.12	.17*	.19**	.12	1

* $p < .05$. ** $p < .01$.

Box 7
Statistically Significant Correlations Between Standardized Test Scores and Text-Based Writing Scores

Nelson-Denny Comprehension raw score and:

- Essay: Proportion functional elements of total elements: $r = .14$*
- Essay: Persuasive quality score: $r = .16$*
- Essay: Word count: $r = .17$*
- Summary: Proportion of main ideas from source text: $r = .29$**
- Summary: Quality score: $r = .31$**
- Summary: Percentage of academic words: $r = .15$*

WJ III Writing Fluency raw score and:

- Essay: Proportion functional elements of total elements: $r = .23$**
- Essay: Persuasive quality score: $r = .28$**
- Essay: Word count: $r = .20$**
- Summary: Quality score: $r = .14$*

* $p < .05$. ** $p < .01$.

Relationship between self-efficacy ratings and teacher judgments and text-based writing measures. Intercorrelations between self-efficacy ratings, teacher judgments, and the text-based writing measures are shown in Table 14, and the statistically significant correlations are listed in Box 8. The intercorrelations between the text-based writing measures are slightly different in Tables 13 and 14 because only students with complete self-efficacy and teacher judgment data were included in the data summarized in Table 14. As with the standardized measures, the significant correlations between self-efficacy ratings and teacher judgments and text-based writing ability were moderate.

Table 14
Intercorrelations Between Self-Efficacy Ratings, Teacher Judgments, and Text-Based Writing Scores

Variable	1	2	3	4	5	6	7	8	9	10
1. Self-efficacy	1									
2. Teacher judgments	.30**	1								
3. Essay: Proportion functional elements	.24**	.22**	1							
4. Essay: Persuasive quality score	.08	.17*	.19*	1						
5. Essay: Word count	−.09	.10	−.07	.47**	1					
6. Essay: % academic words	.10	.11	.06	−.01	.04	1				
7. Summary: Proportion of main ideas	.09	.34**	.08	.08	.15	.03	1			
8. Summary: Quality score	.22**	.33**	.09	.17*	.24**	.10	.62**	1		
9. Summary: Word count	.01	.23**	−.08	.15	.37**	.07	.46**	.54**	1	
10. Summary: % academic words	.02	.23**	.11	−.03	−.01	.09	.14*	.15	.05	1

* $p < .05$. ** $p < .01$.

Box 8
Statistically Significant Correlations Between Self-Efficacy Ratings, Teacher Judgments, and Text-Based Writing Scores

Self-efficacy ratings and:

- Essay: Proportion functional elements of total elements: $r = .24$**
- Summary: Quality score: $r = .22$**

Teacher judgments and:

- Essay: Proportion functional elements of total elements: $r = .22$**
- Essay: Persuasive quality score: $r = .17$*
- Summary: Proportion of main ideas from source text: $r = .34$**
- Summary: Quality score: $r = .33$**
- Summary: Word count: $r = .23$**
- Summary: Percentage of academic words: $r = .23$**

* $p < .05$. ** $p < .01$.

Contribution of standardized test scores, self-efficacy ratings, and teacher judgments to text-based writing. A series of hierarchical regressions was completed in order to determine the contribution of general reading and writing ability (as measured by standardized test scores), self-efficacy ratings, and teacher judgments to scores on the persuasive essays and written summaries. We modeled six variables: the proportion of functional persuasive elements in the essay, essay quality, the percentage of academic words in the essay, the proportion of main ideas in the summary, summary quality, and the percentage of academic words in the summary.

College was entered in the first block as a control variable, since there were statistically significant differences in the standardized test scores (Nelson-Denny Reading Comprehension and WJ III Writing Fluency) between the two sites. The standardized scores were entered in the second block, and the third block consisted of the self-efficacy ratings and teacher judgments. The results of the third block indicate whether self-efficacy ratings and teacher judgments were statistically significant predictors of text-based writing skills over and above the contributions of the standardized test scores. The results of the analyses are shown in Tables 15–20. Total R^2, or the total amount of variance accounted for, ranged from 9 percent (percentage of academic words in the summary) to 18 percent (percentage of academic words in the essay and summary quality).

Proportion of functional elements in the essay. As seen in Table 15, in the full model for the proportion of functional elements in the essay (Model 3, taking into account the standardized test scores, self-efficacy ratings, and teacher judgments), the change in R^2 was statistically significant ($F = 4.580$, $p = .012$), which can be explained by the contributions of the WJ III Writing Fluency scores and the self-efficacy ratings. The change in R^2 in Model 2 (standardized test scores only, without self-efficacy ratings or teacher judgments) was also significant ($F = 7.106$, $p = .001$), which can be attributed to the contribution of the standardized writing scores.

Table 15
Proportion of Functional Elements in the Essay: Summary of Hierarchical Regression Analysis

	Model 1				Model 2				Model 3			
Variable	***B***	***SE B***	**β**	***p***	***B***	***SE B***	**β**	***p***	***B***	***SE B***	**β**	***p***
Constant	.532	.027	—	.000	.200	.101	—	.031	−.203	.162	—	.212
College	−.008	.043	−.015	.850	−.002	.042	−.004	.956	−.019	.041	−.035	.647
Nelson-Denny Comprehension					.001	.002	.059	.473	.001	.002	−.025	.768
WJ III Writing Fluency					.013	.004	.266	.001	.12	.004	.245	.002
Self-efficacy									.004	.002	.173	.034
Teacher judgment									.002	.002	.129	.114
R^2 Δ		.000				.083				.051		
F for change in R^2		.036		.850		7.106		.001		4.580		.012

Note. College was represented as a dummy variable with College 1 as the reference (0) and College 2 coded as 1. Total R^2 = .135.

Essay quality. The change in R^2 in the full model for essay quality was not statistically significant, as seen in Table 16, although, again, the WJ III Writing Fluency scores made a significant contribution. However, the change in R^2 in Model 2 (standardized scores only) was significant ($F = 4.960$, $p = .008$), which was a result of the contribution of the WJ III Writing Fluency scores.

Table 16
Essay Quality: Summary of Hierarchical Regression Analysis

	Model 1				Model 2				Model 3			
Variable	***B***	***SE B***	**β**	***p***	***B***	***SE B***	**β**	***p***	***B***	***SE B***	**β**	***p***
Constant	2.673	.078	—	.000	1.852	.274	—	.000	1.364	.489	—	.006
College	−.238	.126	−.149	.060	−.224	.124	−.140	.073	−.244	.125	−.152	.053
Nelson-Denny Comprehension					.003	.005	.048	.554	.000	.006	.006	.947
WJ III Writing Fluency					.032	.012	.223	.006	.31	.012	.212	.010
Self-efficacy									.001	.005	.021	.802
Teacher judgment									.007	.005	.127	.130
R^2 Δ		.022				.058				.016		
F for change in R^2		3.585		.060		4.960		.008		1.351		.262

Note. College was represented as a dummy variable with College 1 as the reference (0) and College 2 coded as 1. Total R^2 = .096.

Percentage of academic words in the essay. Neither Model 2 nor Model 3 was statistically significant in explaining the percentage of academic words used in the essay, as shown in Table 17, suggesting that performance could not be predicted by the standardized test scores, self-efficacy ratings, or teacher judgments.

Table 17
Percentage of Academic Words in the Essay: Summary of Hierarchical Regression Analysis

	Model 1				Model 2				Model 3			
Variable	***B***	***SE B***	**β**	***p***	***B***	***SE B***	**β**	***p***	***B***	***SE B***	**β**	***p***
Constant	5.117	.259	—	.000	4.874	.934	—	.000	2.312	1.665	—	.167
College	−2.226	.418	−.389	.000	−2.140	.423	−.374	.000	−2.246	.425	−.393	.000
Nelson-Denny Comprehension					.023	.018	.100	.199	.011	.019	.048	.558
WJ III Writing Fluency					−.021	.040	-.040	.606	−.027	.040	−.052	.495
Self-efficacy									.022	.019	.093	.236
Teacher judgment									.018	.016	.093	.237
R^2 Δ		.152				.009				.019		
F for change in R^2		28.416		.000		.842		.433		1.816		.166

Note. College was represented as a dummy variable with College 1 as the reference (0) and College 2 coded as 1. Total R^2 = .180.

Proportion of main ideas from the source text in the summary. As seen in Table 18, the change in R^2 for the proportion of main ideas in the summary was statistically significant in the full model ($F = 7.659$, $p = .001$). This result was explained by the significant contributions of the Nelson-Denny Comprehension scores and the teacher judgments. Model 2 was also significant ($F = 5.190$, $p = .007$) because of a significant contribution of the standardized reading scores.

Table 18
Proportion of Main Ideas in the Summary: Summary of Hierarchical Regression Analysis

	Model 1				Model 2				Model 3			
Variable	***B***	***SE B***	**β**	***p***	***B***	***SE B***	**β**	***p***	***B***	***SE B***	**β**	***p***
Constant	.242	.018	—	.000	.146	.063	—	.023	−.031	.109	—	.777
College	−.022	.029	−.060	.451	−.008	.029	−.021	.790	−.015	.028	−.042	.584
Nelson-Denny Comprehension					.004	.001	.259	.002	.003	.001	.181	.032
WJ III Writing Fluency					−.001	.003	−.027	.735	−.002	-.050	-.010	.527
Self-efficacy									−.001	.001	−.044	.584
Teacher judgment									.004	.001	.313	.000
R^2 Δ		.004				.062				.084		
F for change in R^2		.571		.451		5.190		.007		7.659		.001

Note. College was represented as a dummy variable with College 1 as the reference (0) and College 2 coded as 1. Total $R^2 = .149$.

Summary quality. The change in R^2 was statistically significant in the full model for summary quality ($F = 8.725$, $p = .000$) as well as Model 2 ($F = 5.178$, $p = .007$). In the full model, as seen in Table 19, only the teacher judgments made a significant contribution to summary quality, and in Model 2, only the standardized reading scores made a significant contribution.

Table 19
Summary Quality: Summary of Hierarchical Regression Analysis

	Model 1				Model 2				Model 3			
Variable	***B***	***SE B***	**β**	***p***	***B***	***SE B***	**β**	***p***	***B***	***SE B***	**β**	***p***
Constant	8.333	.278	—	.000	6.357	.977	—	.000	1.385	1.670	—	.408
College	−.962	.448	−.168	.033	−.753	.442	−.131	.091	−.960	.426	−.167	.026
Nelson-Denny Comprehension					.056	.019	.240	.004	.030	.019	.129	.117
WJ III Writing Fluency					.012	.041	.024	.764	−.002	.040	−.004	.956
Self-efficacy									.028	.018	.118	.133
Teacher judgment									.054	.016	.272	.001
R^2 Δ		.028				.060				.092		
F for change in R^2		4.607		.033		5.178		.007		8.725		.000

Note. College was represented as a dummy variable with College 1 as the reference (0) and College 2 coded as 1. Total $R^2 = .180$.

Percentage of academic words in the summary. As shown in Table 20, the full model of the percentage of academic words in the summary was statistically significant ($F = 3.901$, $p = .022$) because of the significant contribution of teacher judgments.

Table 20
Percentage of Academic Words in the Summary: Summary of Hierarchical Regression Analysis

	Model 1				Model 2				Model 3			
Variable	***B***	***SE B***	**β**	***p***	***B***	***SE B***	**β**	***p***	***B***	***SE B***	**β**	***p***
Constant	4.982	.313	—	.000	4.241	1.125	—	.000	2.271	1.980	—	.253
College	−1.073	.505	−.166	.035	−.933	.510	−.144	.069	−1.019	.505	−.158	.045
Nelson-Denny Comprehension					.038	.022	.145	.083	.024	.023	.092	.290
WJ III Writing Fluency					.018	.048	−.031	.705	−.027	.047	−.046	.567
Self-efficacy									−.013	.022	−.050	.544
Teacher judgment									.051	.018	.231	.006
R^2 Δ		.028				.019				.046		
F for change in R^2		4.513		.035		1.536		.218		3.901		.022

Note. College was represented as a dummy variable with College 1 as the reference (0) and College 2 coded as 1. Total $R^2 = .092$.

6.5 Student Retrospective Reports

Retrospective reports were obtained in individualized interviews with 28 students, selected on the basis of their interest and availability. The interviews focused on one of the text-based writing tasks, the summary of the newspaper article on intergenerational conflict in the workplace. Writing samples produced by eight of the students are provided in Appendix A. Through the retrospective reports, we sought to learn about students' understanding of the task's instructions and obtain first-person descriptions of strategies they used to read and summarize the article. An analysis of the interviews indicated that all of the interviewees understood the information in the source text. However, many were not able to define "summary" with precision.

The central demand of text-based summarization is to identify the most important information in the source text and then paraphrase it in writing so that it captures the gist. Although it would be inappropriate to include a personal opinion or information extraneous to that found in the source text in a summary of a text, some students stated that they expressed their opinion of the material. The following quotations from the interviews illustrate students' conceptualization of a summary (the last two showing examples of confusion regarding what constitutes a summary).

> [A summary is] just an overall explanation of what the article is about or what the book is about and the main details that are in it that I took from it and that I feel other people should take out of that and that are important in the article.

> Summary basically means the main idea, nobody is going to want to read a five-page paper, they want the basics. Just tell me what happens, give me the main points and the main ideas so I don't have to read the whole thing. It just tells me what I need to know.

> When you summarize a paper, it's reading through it, jotting down notes or … annotating, highlighting areas that is important.

> Summary meant to me like somewhat where they was talking about in the paragraph and kind of plot the main things, make sure you get the main details out of it.

> I was kind of confused when they said write … a summary on a article because I mean it was kind of … it wasn't like facts, it was kind of like a story kind of, so I couldn't understand it, but I guess it was more opinion on what you got out of it.

> Like, you read something, then in your own words—not write a whole page, but a very detailed information about what you read in your own words.

While many students could not precisely describe what a summary is, the strategies they reported suggested a good level of implicit knowledge of the task's requirements. Students reported the use of a variety of strategies to identify the main

ideas for inclusion in the summary. Their reports of these strategies suggested a good understanding of the drafting process in writing, as illustrated in the following quotations from our interviews.

Reading and annotating:

> … read it first and probably annotate it. Then go back and look at the things I underlined to write the summary. [Annotation means to] ask questions and underline stuff and things like that. … First I glanced through it and looked at what was on it, and then I read it carefully and underlined important things. … Not reading the actual thing more than once, just because once I annotated it, I wouldn't have to read it, after highlighting stuff, not highlighting but underlining and stuff. I didn't have to read it again, because I underlined the important information.

Highlighting and annotating the text:

> When I'm reading it, I highlight and annotate the important information because I knew I was writing a summary, so then what I highlighted was what I was going to write in the summary. … When I highlighted and annotated it, I use that as my paraphrasing it … Our teacher taught us … instead of just reading an article and just then trying to write about it, as you go through, try to pinpoint the summaries and the important information, so then when you go back to writing, you are not like, oh, so what did they say in the paragraph? So you are a little bit more organized with your work. … I took the information that I highlighted and just paraphrased it and put it in my own words.
>
> Every time you underline it, so then when you want to write out like what you were talking about, you could just skim and say, I can take that from that and put it in a little summary, and then I'll know exactly what to say.
>
> I read one thing at a time, and if I see anything that I think is important, I highlight it, or I make a note about it. … I highlight anything that jumps out at me that I think is important.
>
> I annotated the article, and it said to summarize it, so I pretty much read it, annotated, summarized what I marked when I read it. Went back and pretty much filled in the

> blanks from what I annotated and just put it in my own words. … Annotating is like, you can either highlight or you can either use a pen, mark what is important to you. Our teacher usually teach us to write inside the, um, right in here, in the margins of, you know, each line here. … I annotate it and highlight it, mark it so that I could just go back and look, and I don't have to, you know, keep reading. You just skim through this, what you got.

> I always annotate everything I read because it helps me understand it better. So I pick out the little things to help me write what I have to write, so it gives me little ideas, so I can go back to it and read my ideas and write it in the essay, I guess. … Annotate is when you like pick out like a few sentence … or like little words in the sentences and write to the side, that's how I've always been taught to write, to the side, or like pick from the paragraph or sentence and write to the side … the important parts or something that stood out to me, or what I thought of the sentence.

Using annotations while writing:

> I read it once, and then I went back, and after I did the annotations, I did read little parts to help me do it again, to help me write it.

Highlighting, making notes:

> I read one thing at a time, and if I see anything that I think is important, I highlight it, or I make a note about it. … I highlight anything that jumps out at me that I think is important.

Underlining text:

> I usually use my pen, go through each paragraph first because sometimes I can get lost and skip lines, so I usually hold my pen and I go through it, and I kind of process what I read. I don't read fast, I'm not one of those people who can read fast and understand it, so I usually read slow and basically go over main stuff, and sometimes when I pick up important things, I'll underline it because that's what I need to put, that's like the main idea, so like aspects I need to add to my summary.

Rereading:

> I read it twice. I read my articles twice. In case I didn't get something the first time, I do it the second time.
>
> I went back when I wanted to start writing my summary. After I read it and I said, okay, this is what I'm going to write about, this is what I'm going to say, then I go back and say, okay, wait, before I can write it down, I need to know what to write. So I need to look in each paragraph to know what are they saying to me on my opinion.

Skimming and rereading while writing:

> When I went to do it, I just went through each paragraph and picked out important things. So I basically read it, and then I skimmed over it again. …You really shouldn't be reading an article once, I believe, I mean, you should read it more than once because you are not going to … everyone gets distracted, you're not going to understand it the first time, so even if you don't understand it, I think you should read it more than once.

Formulating a topic sentence:

> I always write, you know, a little topic sentence. That gives me an idea of what I'm going to be writing about. … I went back to the question and pretty much put the question into the sentence, and that gives me an idea of what I'm going at.

Selecting information for the summary:

> The main points is what I underlined so that when next I come—I will not forget. I will just see what are the main things and then try to summarize it in my own words. … Because like the main points are there, but there are things that are supporting them. So I just underlined them and just leave the rest. And then use my own words to summarize everything.
>
> I have to break down paragraph by paragraph and put in what I understand in order to write a summary about it. So I kind of pulled out main details from out of each paragraph, and like the small paragraphs I kind of put together and I wrote about that way.

Skimming, underlining, use of prior knowledge:

> Well, I skimmed the article and took some notes and underlined important details that I thought were important. And I related it back to where I work at, and if I was higher up than my other coworkers would they get mad or not. So that's how I envisioned it. … I kind of relate to that because I'm a young employee and I have older employees, too. So I kind of related it back. … not like trying to put personal experiences in there, but try to relate it back to the other coworkers and employer.

Previewing text:

> I always scan through my stuff first, so I kind of scanned through to see what I'll be reading about, and I picked out the word that I thought was difficult to me even though it may seem simple to anybody else was multi-general workplace, and I kind of brought that down within myself so I can understand what I was reading, and then I read paragraph by paragraph and summed it up as I go to paragraph, so by the time I got to the end, I understand what I was reading or so.

Although most students described appropriate strategies, some may have been using strategies that were not optimal for the summarization. For example, a student described a compare–contrast strategy:

> The first thing I did when I started writing was try to show the reader that basically I'm making a comparison, you know. So, that's why I started out with young versus older generations to let them know that, okay, this paragraph is going to be about diversity and the different steps how young people do things and how older people do something.

Although planning and revising are important in writing (Kiuhara, O'Neill, Hawken, & Graham, 2012; Tillema, van den Bergh, Rijlaarsdam, & Sanders, 2011), when students did plan their summaries, the planning was done mentally rather than with the use of explicit organizers, such as diagrams, which are recommended in the literature (Robinson & Kiewra, 1995; Westby et al., 2010). Further, there was little description of meaningful revision in the retrospective reports, and a few students mentioned that their developmental teacher had not asked them to revise their writing. Thus, while an understanding of the nature of drafting was generally well developed, students appeared to lack planning and revision strategies.

7. Discussion

Although there is no disagreement with claims that a large proportion of college entrants in the United States are underprepared for postsecondary academic demands (ACT, 2014; Porter & Polikoff, 2012), there is very limited literature on the actual literacy skills of underprepared students. Recent studies have described college reading and writing requirements (Armstrong et al., 2015; National Center on Education and the Economy, 2013), but there is a shortage of analyses of the skills of specific types of students. The current study fills this gap by providing detailed information on the reading and writing skills and self-efficacy of a sample of developmental reading and writing students. Such information can deepen the understanding of commonly used metrics, such as scores on college placement tests, developmental referrals, grade point average, and academic persistence.

In the current study, we asked to what extent our participants, who were attending developmental education courses in which reading and English were integrated, were able to perform two tasks involving reading and writing that are central to postsecondary learning—text-based summarization and persuasive writing. We also investigated the contribution of standardized test scores, student self-efficacy ratings, and teacher judgments to these skills. The work was conducted by a researcher-practitioner partnership between CCRC and two community colleges in a southern state.

Our key findings are as follows:

1. The researcher-practitioner partnership was successful in developing and implementing an assessment that had direct implications for classroom instruction.
2. The participants still had quite a way to go in order to be ready for college reading and writing, as indicated by both the standardized tests of general reading and writing ability and the project measures, which reflected literacy demands of introductory college courses. The data strongly suggest that the students, even at the top level of the integrated reading and writing course, continued to be underprepared for college reading and writing.

3. There were differences in students' standardized reading scores between the two colleges participating in the study but no differences in standardized writing scores. Students at the two colleges differed in their scores on three of eight text-based writing variables.

4. There was no difference in the standardized reading or writing scores between students in the intermediate and top-level developmental reading and English courses at either college. Students' scores on some of the text-based measures did differ by developmental level, but only at one of the colleges.

5. Both self-efficacy ratings and teacher judgments were relatively high, in contrast with students' relatively low reading and writing scores. However, although self-efficacy ratings and teacher judgments were high, the correlation between these two measures was moderate, suggesting that for individual students, there were discrepancies between students' and teachers' levels of confidence in their proficiency.

6. Correlations between the standardized measures and the project-developed text-based writing measures, where statistically significant, were moderate, suggesting weak predictive relationships. Of the 45 correlations we ran, only four exceeded .40.

7. A series of hierarchical regressions suggests the importance of general reading skills for text-based summarization, and general writing skills for text-based persuasive essay writing. Although both tasks required both reading and writing, performance on the two tasks was explained by different skills.

8. Students' retrospective reports provided an interesting and useful window into their text-based writing skills. Although students had difficulty describing the nature of text-based summarization, they engaged in a variety of appropriate strategies to perform the task.

In the remainder of this section, we discuss the implications of our findings.

7.1 Partnership Development

The researcher-practitioner partnership was successful in that communication among actors was consistent and substantive from the beginning and the assessment was completed as planned. The faculty at the participating sites were instrumental in the design of the research tasks as well as the procedures for data collection. A few changes were made to procedures based on piloting and on collaborative discussion between the partners. Careful coordination by both researchers and college staff, along with regular phone conferences, sharing of material, and an in-person one-day retreat, seem to have been key elements in the success of the collaboration.

Over an extended period of time involving in-depth work at both institutions and with lead practitioners, a few strategies emerged as important points to consider for future researcher-practitioner partnerships. First, it is important for the researcher to thoroughly understand the contextual factors within the practitioner institutions. These factors may include the classroom environment, curriculum, and instructional approaches, and challenges faced by students. Researchers also benefit from feedback from instructors or other practitioners who form the partnership on research questions and data collection procedures. As part of the collaboration, researchers can learn about contextual elements from their practitioner partners and better understand both what issues to study and how to most effectively study the area of focus. For example, the feedback we received from faculty about how to phrase prompts helped ensure that the data we collected reflected the students' actual skills and were not biased by prompts that were presented in a way unfamiliar to students.

Second, the researcher-practitioner partnership's emphasis on mutualism requires an established relationship of rapport and trust. This relationship is an important factor in gaining the access necessary to carry out research activities, in addition to maintaining an open communication setting in which sincere feedback is welcomed. CCRC's partnership with the two community colleges was grounded in a relationship that initially began with their participation in fieldwork for CCRC's broader study of the statewide developmental course redesign. These site visits included classroom observations, interviews, and student focus groups with our lead faculty partners' classes on multiple occasions. By engaging in these activities prior to forming the partnership that is highlighted in the current study, lead

faculty came to be familiar with CCRC researchers and our work. A sense of familiarity and trust characterized subsequent interactions between the partners. The collegial atmosphere allowed us to have frank conversations about logistical issues, such as the length of student testing, which ultimately ensured a smooth administration of the assessment.

Finally, the partnership experience highlights the importance of maintaining flexibility in the research design and being able to respond to logistical challenges raised by practitioners. For instance, as stated previously, the practitioner partners raised concerns regarding the length of the assessments. Recruitment challenges experienced during the pilot administration corroborated the faculty's concerns. In response, we reviewed our tasks and, using insights gained from our pilot data, revised our assessment in a way that alleviated the faculty's concerns while maintaining the scholarly integrity of the assessment. Through this process, we not only maintained positive relationships with our partners but also maximized our likelihood for a high yield in terms of student recruitment.

7.2 Students' Readiness for College-Level Literacy Tasks

A major goal of this study was to determine how close developmental reading and English students were to being ready for introductory college-level text-based writing tasks. Based on literature detailing the literacy demands of college-level work (Carson et al., 1992; J. M. Jackson, 2009; McAlexander, 2003; O'Neill et al., 2012; Yancey, 2009), the scores on both the standardized and project-developed measures used in the current research suggest that the participants remained underprepared for the reading and writing demands of their upcoming college-credit courses.

The participants tested at the lower end of the average range for end-of-year 12th graders on the standardized reading and writing measures (22nd and 27th percentiles, respectively). The scores can be compared with those found in previous research on developmental education students (MacArthur, Philippakos, & Graham, 2015; Perin et al., 2013). In the Perin et al. (2013) study, upper level developmental education students obtained a mean raw score of 30.95 (*SD* = 15.22), or 41 percent correct, on the Nelson-Denny Comprehension subtest, which is similar to the score of 29.32 (*SD* = 12.29), or 38 percent correct, found in the current sample of intermediate and upper level developmental education students. MacArthur et al. (2015) reported a mean raw score of 19.9 (*SD* = 4.5), or 50 percent correct, for upper level developmental students on the WJ III Writing

Fluency subtest, which is somewhat similar to the mean raw score of 22.50 (*SD* = 5.25), or 56 percent correct, in the present sample.

The text-based tasks were used to assess readiness for college-level literacy demands because they required both reading comprehension and writing skill and were typical of classroom assignments. Across the whole sample, when summarizing a newspaper article, the participants included 19 percent of the main ideas from the source text. Although it is not expected that all of the main ideas would be included in a summary written even by the most proficient writer, 19 percent of the main ideas seems low if the summary is to capture the gist of the source text. Further, although norms are not available for this task, performance for this sample fell below that of two other developmental education samples, 28 percent in Perin et al. (2003) and 42 percent in Perin et al. (2013). The quality of the written summaries, measured by an analytic rubric that focused on four components of summarization, also tended to be somewhat low, with a mean score of 8.09 (*SD* = 2.64) on a 16-point scale.

Students also demonstrated weakness in text-based persuasive essay writing, with a mean score of 2.58 (*SD* = .80) on a 7-point holistic scale. Across students, almost one half of the content written, measured in terms of functional persuasive elements, was not helpful in the development of a persuasive argument. There are no prior studies of the use of academic words in the writing of adolescents or adults, but one study found that 1 percent of the words used in the writing of typically developing fifth graders were academic words (Olinghouse & Wilson, 2013). The fact that academic words accounted for only 3 percent of the words written by the college students assessed in the current study may point to the need to develop this skill further.

7.3 Differences by College and Developmental Level

The differences in reading scores between the two colleges in the study is a reminder of the variation that exists across institutions and the dangers of generalizing from one sample to another. Although it was beyond the scope of the current study to investigate college-based differences that might account for the differing reading scores, future studies could test hypotheses about the relation of variables such as students' background, curriculum, and teachers' pedagogical styles that might explain differences in skills across institutions.

Although it would not be expected that students at the intermediate developmental level would be fully ready for college literacy demands, an unexpected finding of this study is the similarity in level of skill between the students attending intermediate and top-level courses. Scores on the standardized reading and writing tests were similar between the two levels in both colleges. The majority of comparisons between levels suggested that students at the two levels were more similar than different in their reading and writing skills. It was only at College 2, which had been integrating reading and writing instruction for a number of years, that differences were found, and these differences were only on four of the eight text-based measures (all with top-level students outperforming intermediate-level students); there were no differences on either of the standardized measures. In College 1, which had just begun integrating its developmental reading and writing instruction, there were no differences in the standardized test scores or on any of the text-based writing measures as a function of developmental course level. It is possible that placement procedures at this particular college were resulting in some inaccurate course referrals.

7.4 Student Self-Efficacy and Teacher Judgments

This assessment study also investigated the levels of confidence that students had in their reading and writing ability, and their instructors' judgments of that ability. Self-efficacy was surprisingly high, given students' low reading and writing skills, with a mean of 81.60 (*SD* = 11.43) out of 100 points. This mean score is higher than the mean of 70.3 (*SD* = 14.7), also on a scale of 100, reported for top-level developmental education students by MacArthur, Philippakos, and Ianetta (2015). The mean teacher judgment score of 71.84 (*SD* = 14.17) in the current data was also unexpectedly high. Thus, although the teacher judgments were lower than the student ratings, both seemed inflated in the context of students' reading and writing scores. Future efforts could focus on familiarizing both students and instructors with the reading and writing demands of introductory college-level courses. Although developmental reading and writing courses often focus on preparation for the first level of college English, it would be beneficial to students if a wider view of literacy readiness were adopted, incorporating the reading and writing demands of disciplinary courses as well.

7.5 Relationships Between Measures

Another question in this study concerned the relationships between the assessment tasks. Students' performance scores on the two text-based writing tasks had a statistically significant but relatively weak relationship to each other ($r = .16$, $p < .05$ for summary and essay quality), suggesting that the two tasks called for different skills. This possibility is supported by the different relationships of the standardized reading and writing scores to the two tasks. Standardized reading scores had a higher correlation with summary quality ($r = .31$, $p < .01$) than with essay quality ($r = .16$, $p < .05$), and standardized writing scores had a higher correlation with essay quality ($r = .28$, $p < .01$) than with summary quality ($r = .14$, $p < .05$). Although both tasks required both reading and writing, it appears that reading skills were more important for the written summarization task and writing skills were more important for the text-based persuasive essay.

Another notable finding was that, although the standardized measures were reliably related to performance on the text-based writing measures, even the highest correlations were moderate, suggesting that they were tapping different skill sets. Of the 45 correlations we ran between the scores on reading and writing skills, only five exceeded $r = .40$: essay quality and essay word count ($r = .51$, $p < .01$), essay quality and the percentage of academic words in the essay ($r = .57$, $p < .01$), proportion of main ideas in the summary and summary quality ($r = .61$, $p < .01$), proportion of main ideas in the summary and summary word count ($r = .47$, $p < .01$), and summary quality and summary word count ($r = .50$, $p < .01$). Of these five relatively strong correlations, three concerned word count. Although the direction of the relationship between each pair of variables is unknown, a hypothesis can be proposed that working with students to lengthen their writing samples may help them improve their writing.

The self-efficacy ratings were significantly correlated with teacher judgments, but the correlation was only moderate ($r = .30$, $p < .01$), suggesting that for individual students, there were discrepancies between students' and teachers' level of confidence in the student's proficiency.

Self-efficacy ratings correlated significantly with only two of the eight text-based writing variables, whereas teacher judgments correlated significantly with six of these variables. Therefore, the current data suggest that the teachers may be better than their

students at predicting the students' reading and writing skills. Although previous research has found self-efficacy to be a significant predictor of literacy skills (MacArthur, Philippakos, & Graham, 2015; Pajares & Valiante, 2006; Proctor et al., 2014), the current study appears to be unique in being able to compare self-efficacy ratings with teacher judgments for the same sample.

7.6 Contribution of Standardized Test Scores, Self-Efficacy Ratings, and Teacher Judgments to Variance in Text-Based Writing

We also examined the relative contribution of standardized test scores, self-efficacy ratings, and teacher judgments to the text-based writing measures, controlling for college attended. Results of hierarchical regression analyses indicated the importance of the standardized writing scores in predicting the proportion of functional elements in the persuasive essay and the quality of the essay, while the standardized reading scores were important in predicting the proportion of main ideas in the summary and the quality of the summary. Thus, as suggested by the correlations between measures, improvement in text-based summarization may require particular attention to reading comprehension skills, while improvement in text-based persuasive essay writing may depend more on developing general writing skills.

Self-efficacy was only important in predicting the proportion of functional elements in the essay, and teacher judgments were only important in predicting the percentage of academic words in the summary. Thus, self-efficacy and teacher judgments had only a small explanatory role in text-based writing once the specific college and standardized scores were taken into account. Previous research suggests that self-efficacy is a reliable predictor of literacy performance (e.g., MacArthur, Philippakos, & Graham, 2015; Martinez et al., 2011) and that teacher judgments correlate with literacy skills, although relationships are stronger at higher skill levels (Begeny et al., 2011; Feinberg & Shapiro, 2009). Perhaps the lack of high achievers in the current sample, in conjunction with the ceiling effect on self-efficacy (with many scores at the highest part of the scale), at least partly explains the relatively small amount of variance of self-efficacy and teacher judgments, compared with standardized test measures, in accounting for the text-based writing scores.

7.7 Students' Task Understanding as Revealed by Retrospective Reports

Our last research question asked how students described task demands and the strategies they used to write a text-based summary. The retrospective reports suggested an uneven ability to articulate what was required to summarize text. However, most of the students interviewed described summarization strategies that reflected a good understanding of the nature of summarization (representing the gist, or the main ideas of a source text) even though they could not define summarization precisely. The strategies students used to summarize the newspaper article included previewing the source text, carefully selecting the important ideas from the article, reading, rereading, skimming, annotating the text, highlighting and underlining important information in the article, and applying prior knowledge to understand the content. Future research could compare students' reports of their strategies with observations of the described strategies in action in order to learn to what extent and how proficiently students applied the strategies they mentioned in their retrospective reports.

7.8 Motivation and Effort

Since motivation and effort are important variables in adult learning (Liu, Bridgeman, & Adler, 2012; MacArthur, Philippakos, & Graham, 2015; Mellard, Krieshok, Fall, & Woods, 2013; Merriam & Bierema, 2013), it is worth noting the current students' self-reports on these variables. On the student background questionnaire, the participants reported moderate to strong motivation and effort on the assessment. Previous research suggests that self-reports of motivation are predictive of performance on reading tests (Retelsdorf, Köller, & Möller, 2011); however, there was a ceiling effect in our sample (most ratings were near the top of the scale), ruling out the opportunity of detecting possible effects. Although not a primary concern of the current research, in view of the importance of these variables, along with anecdotal reports from instructors that some students were not applying themselves as well as they could, the current research could be extended in the future by more robust measures of motivation and effort. Moreover, motivation and effort may interact with specific academic skills, such as those measured in this study in predicting literacy performance in introductory college-level courses.

7.9 Implications for Instruction and Placement Policy

Our ability to draw implications for developmental education is constrained by the fact that the current study involved only two colleges. Further research is needed to investigate whether the current study's findings can be generalized to developmental education students across the country. However, pending such findings, tentative implications can be drawn.

The low skills of our sample of students on two typical college literacy tasks suggest the need to investigate alternative approaches to curriculum design and/or instruction. For example, prior research with adolescent and adult students suggests that explicit instruction in carefully orchestrated strategies results in better reading and writing performance (Cook & Bennett, 2014; De La Paz & Felton, 2010; MacArthur & Lembo, 2009; MacArthur, Philippakos, & Ianetta, 2015; Reynolds & Perin, 2009; Simpson, 1986). Contextualizing strategy instruction in subject matter students are learning in discipline-area classrooms may enhance its effects (Perin, 2011; Shanahan, Shanahan, & Misischia, 2011). Strategy instruction, especially when contextualized, differs from "business as usual" in college developmental classrooms (Grubb & Gabriner, 2013b) and would require well-planned professional development.

In addition, placement policy is a serious and ongoing concern of community colleges, which serve high numbers of underprepared students (Hassel & Giordano, 2015; Jaggars, Hodara, Cho, & Xu, 2015). The finding that students' developmental course level was associated with only a few of the variables measured in the current assessment suggests a need to review placement policy, at least in the two participating colleges. Future research could determine whether this is a merely a local finding or whether the lack of difference between developmental levels is more widespread. Such a finding would inform policy discussions about developmental education, including the question of how many levels of developmental courses should be offered.

Another concern is the effectiveness of developmental education in preparing students for the literacy demands of their introductory disciplinary courses. The current sample appeared to demonstrate a need for ongoing academic support after exiting developmental education if they were to succeed at the college level.

8. Conclusion

This study offers a method of understanding the educational needs of low-skilled postsecondary students that is deeper than that permitted by the traditional assessment metrics of course completion, grades, persistence, and degree award. The use of literacy tasks that are typical of introductory college-level coursework provides insight into the skills that students need to be ready to read and write at the college level. Although the low skills revealed in the study may be disappointing to the students, instructors, and administrators who devote a significant amount of time and effort to developmental education, awareness of the reality of the situation is the first step toward positive change.

The current assessment procedure appears to have good potential for adding to more general measures of college readiness, such as college placement tests. Limitations have been identified in the use of single measures for placing students into developmental education (Hughes & Scott-Clayton, 2011), and it is possible that the use of multiple measures might lead to more accurate course placement (Scott-Clayton, 2012). The battery used in the current study, with its combination of standardized and project-developed measures and its inclusion of self-efficacy ratings, teacher judgments, and student retrospective reports, may provide a framework for the future development of multiple assessment measures that are educationally meaningful. Predictive, longitudinal data would be needed to compare outcomes for such a battery, both with other sets of multiple measures and with the traditional single measures. Finally, an essential component of this study was the researcher-practitioner partnership. The insights and practical concerns of the college personnel made it possible for the team to develop an assessment approach that contained meaningful tasks and obtained a multifaceted perspective on student abilities.

References

Acker, S. R., & Halasek, K. (2008). Preparing high school students for college-level writing: Using ePortfolio to support a successful transition. *Journal of General Education, 57*(1), 1–15. doi:10.1353/jge.0.0012

ACT. (2014). *The condition of college and career readiness 2014: National*. Retrieved from http://www.act.org/research/policymakers/cccr14/index.html

Aghaie, R., & Zhang, L. J. (2012). Effects of explicit instruction in cognitive and metacognitive reading strategies on Iranian EFL students' reading performance and strategy transfer. *Instructional Science*, *40*(6), 1063–1081. doi:10.1007/s11251-011-9202-5

Armstrong, S. L., Stahl, N. A., & Kantner, M. J. (2015). *What constitutes 'college-ready' for reading? An investigation of academic text readiness at one community college* (Technical Report No. 1). Retrieved from Center for the Interdisciplinary Study of Literacy and Language website: http://www.niu.edu/cisll/documents/index.shtml

Atkinson, T. S., Zhang, G., Phillips, S. F., & Zeller, N. (2014). Using word study instruction with developmental college students. *Journal of Research in Reading, 37*(4), 433–448. doi:10.1111/1467-9817.12015

Bailey, T. R., Jeong, D. W., & Cho, S.-W. (2010). Referral, enrollment, and completion in developmental education sequences in community colleges. *Economics of Education Review*, *29*(2), 255–270.

Barnett, E. A., Bork, R. H., Mayer, A. K., Pretlow, J., Wathington, H. D., & Weiss, M. J. (with Weissman, E., Teres, J., & Zeidenberg, M.). (2012). *Bridging the gap: An impact study of eight developmental summer bridge programs in Texas*. New York, NY: National Center for Postsecondary Research.

Begeny, J. C., Krouse, H. E., Brown, K. G., & Mann, C. M. (2011). Teacher judgments of students' reading abilities across a continuum of rating methods and achievement measures. *School Psychology Review*, *40*(1), 23–38.

Bettinger, E. P., Boatman, A., & Long, B. T. (2013). Student supports: Developmental education and other academic programs. *Future of Children*, *23*(1), 93–115.

Bohn-Gettler, C. M., & Kendeou, P. (2014). The interplay of reader goals, working memory, and text structure during reading. *Contemporary Educational Psychology*, *39*(3), 206–219. doi:10.1016/j.cedpsych.2014.05.003

Bråten, I., Strømsø, H. I., & Britt, M. A. (2009). Trust matters: Examining the role of source evaluation in students' construction of meaning within and across multiple texts. *Reading Research Quarterly*, *44*(1), 6–28.

Bridgeman, B., & Carlson, S. B. (1984). Survey of academic writing tasks. *Written Communication*, *1*(2), 247–280. doi:10.1177/0741088384001002004

Brown, A. L., & Day, J. D. (1983). Macrorules for summarizing texts: The development of expertise. *Journal of Verbal Learning and Verbal Behavior*, *22*(1), 1–14.

Brown, J. I., Fishco, V. V., & Hanna, G. S. (1993). *The Nelson-Denny Reading Test, Forms G and H.* Itasca, IL: Riverside/Houghton-Mifflin.

Bruning, R., Dempsey, M., Kauffman, D. F., McKim, C., & Zumbrunn, S. (2013). Examining dimensions of self-efficacy for writing. *Journal of Educational Psychology*, *105*(1), 25–38. doi:10.1037/a0029692

Cantrell, S. C., Correll, P., Clouse, J., Creech, K., Bridges, S., & Owens, D. (2013). Patterns of self-efficacy among college students in developmental reading. *Journal of College Reading and Learning*, *44*(1), 8–34. doi:10.1080/10790195.2013.10850370

Carson, J. G., Chase, N. D., Gibson, S. U., & Hargrove, M. F. (1992). Literacy demands of the undergraduate curriculum. *Reading Research and Instruction*, *31*(4), 25–50. doi:10.1080/19388079209558094

Caverly, D. C., Nicholson, S. A., & Radcliffe, R. (2004). The effectiveness of strategic reading instruction for college developmental readers. *Journal of College Reading and Learning*, *35*(1), 25–49.

Chou, M.-H. (2013). Strategy use for reading English for general and specific academic purposes in testing and nontesting contexts. *Reading Research Quarterly*, *48*(2), 175–197.

Coburn, C. E., Penuel, W. R., & Geil, K. E. (2013). *Research-practice partnerships: A strategy for leveraging research for educational improvement in school districts*. New York, NY: William T. Grant Foundation.

Cohen, A. M., & Brawer, F. B. (2008). *The American community college* (5th ed.). San Francisco, CA: Jossey-Bass.

Conley, D. T. (2007). *Redefining college readiness*. Eugene, OR: Educational Policy Improvement Center.

Conley, D. T. (2008). Rethinking college readiness. *New Directions for Higher Education*, *2008*(144), 3–13. doi: 10.1002/he.321

Conley, D. T. (2012). *A complete definition of college and career readiness.* Retrieved from Educational Policy Improvement Center website: http://www.epiconline.org/ccr-definition/

Conley, D. T., & French, E. M. (2014). Student ownership of learning as a key component of college readiness. *American Behavioral Scientist*, *58*(8), 1018–1034. doi:10.1177/0002764213515232

Cook, K. B., & Bennett, K. E. (2014). Writing interventions for high school students with disabilities: A review of single-case design studies. *Remedial and Special Education*, *35*(6), 344–355. doi:10.1177/0741932514523140

Corkill, A. (2007). Review of the Nelson-Denny Reading Test, Forms G and H. In K. F. Geisinger, R. A. Spies, J. F. Carlson, & B. S. Plake (Eds.), *Seventeenth mental measurements yearbook*. Lincoln, NE: Buros Institute of Mental Measurement.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*(2), 213–238.

Crammond, J. G. (1998). The uses and complexity of argument structures in expert and student persuasive writing. *Written Communication*, *15*(2), 230–268. doi:10.1177/0741088398015002004

De La Paz, S., & Felton, M. K. (2010). Reading and writing from multiple source documents in history: Effects of strategy instruction with low to average high school writers. *Contemporary Educational Psychology*, *35*(3), 174–192. doi:10.1016/j.cedpsych.2010.03.001

De La Paz, S., Ferretti, R., Wissinger, D., Yee, L., & MacArthur, C. A. (2012). Adolescents' disciplinary use of evidence, argumentative strategies, and organizational structure in writing about historical controversies. *Written Communication*, *29*(4), 412–454. doi:10.1177/0741088312461591

Doolan, S. M. (2014). Comparing language use in the writing of developmental Generation 1.5, L1, and L2 tertiary students. *Written Communication*, *31*(2), 215–247. doi:10.1177/0741088314526352

Edgecombe, N. (2011). *Accelerating the academic achievement of students referred to developmental education* (CCRC Working Paper No. 30, Assessment of Evidence Series). New York, NY: Columbia University, Teachers College, Community College Research Center.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data (revised edition)*. Cambridge, MA: MIT Press.

Fallahi, C. R. (2012). Improving the writing skills of college students. In E. L. Grigorenko, E. Mambrino & D. D. Preiss (Eds.), *Writing: A mosaic of new perspectives* (pp. 209–219). New York, NY: Psychology Press.

Farrington-Flint, L., Coyne, E., Stiller, J., & Heath, E. (2008). Variability in children's early reading strategies. *Educational Psychology*, *28*(6), 643–661. doi:10.1080/01443410802140958

Farrington-Flint, L., & Wood, C. (2007). The role of lexical analogies in beginning reading: Insights from children's self-reports. *Journal of Educational Psychology*, *99*(2), 326–338. doi:10.1037/0022-0663.99.2.326

Feinberg, A. B., & Shapiro, E. S. (2009). Teacher accuracy: An examination of teacher-based judgments of students' reading with differing achievement levels. *Journal of Educational Research*, *102*(6), 453–462. doi:10.3200/JOER.102.6.453-462

Ferretti, R. P., MacArthur, C. A., & Dowdy, N. S. (2000). The effects of an elaborated goal on the persuasive writing of students with learning disabilities and their normally achieving peers. *Journal of Educational Psychology*, *92*(4), 694–702. doi:10.10377//0022:2-0663.92.4.694

Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist*, *35*(1), 39–50. doi:10.1207/S15326985EP3501_5

Gil, L., Bråten, I., Vidal-Abarca, E., & Strømsø, H. I. (2010). Understanding and integrating multiple science texts: Summary tasks are sometimes better than argument tasks. *Reading Psychology*, *31*(1), 30–68. doi:10.1080/02702710902733600

Graham, S. (1999). Handwriting and spelling instruction for students with learning disabilities: A review. *Learning Disability Quarterly*, *22*(2), 78–98. doi:10.2307/1511268

Griva, E., Alevriadou, A., & Semoglou, K. (2012). Reading preferences and strategies employed by primary school students: Gender, socio-cognitive and citizenship issues. *International Education Studies*, *5*(2), 24–35.

Grubb, W. N., & Gabriner, R. (2012). *Basic skills education in community colleges: Inside and outside of classrooms*. New York, NY: Routledge.

Hale, G., Taylor, C., Bridgeman, B., Carson, J., Kroll, B., & Kantor, R. (1996). *A study of writing tasks assigned in academic degree programs* (RR-95-44). Princeton, NJ: Educational Testing Service.

Harrison, G. L., & Beres, D. (2007). The writing strategies of post-secondary students with writing difficulties. *Exceptionality Education Canada*, *17*(2), 221–242.

Hassel, H., & Giordano, J. B. (2015). The blurry borders of college writing: Remediation and the assessment of student readiness. *College English*, *78*(1), 56–80.

Hillocks, G., Jr. (2011). *Teaching argument writing, grades 6–12: Supporting claims with relevant evidence and clear reasoning*. Portsmouth, NH: Heinemann.

Hodara, M., & Jaggars, S. S. (2014). An examination of the impact of accelerating community college students' progression through developmental education. *Journal of Higher Education*, *85*(2), 246–276. doi:10.1353/jhe.2014.0006

Hoge, R. D., & Coladarci, T. (1989). Teacher-based judgments of academic achievement: A review of literature. *Review of Educational Research*, *59*(3), 297–313. doi:10.3102/00346543059003297

Holschuh, J. P. (2014). The common core goes to college: The potential for disciplinary literacy approaches in developmental literacy classes. *Journal of College Reading and Learning*, *45*(1), 85–95. doi:10.1080/10790195.2014.950876

Holschuh, J. P., & Aultman, L. P. (2009). Comprehension development. In R. F. Flippo & D. C. Caverly (Eds.), *Handbook of college reading and study strategy research* (2nd ed., pp. 121–144). New York, NY: Routledge.

Hughes, K. L., & Scott-Clayton, J. (2011). Assessing developmental assessment in community colleges. *Community College Review*, *39*(4), 327–351. doi:10.1177/0091552111426898

Jackson, J., & Kurlaender, M. (2014). College readiness and college completion at broad access four-year institutions. *American Behavioral Scientist*, *58*(8), 947–971. doi:10.1177/0002764213515229

Jackson, J. M. (2009). Reading/writing connection. In R. F. Flippo & D. C. Caverly (Eds.), *Handbook of college reading and study strategy research* (2nd ed., pp. 145–173). New York, NY: Routledge.

Jaggars, S. S., Hodara, M., Cho, S.-W., & Xu, D. (2015). Three accelerated developmental education programs: Features, student outcomes, and implications. *Community College Review*, *43*(1), 3–26. doi:10.1177/0091552114551752

Jenkins, D., & Boswell, K. (2002). *State policies on community college remedial education: Findings from a national survey* (Technical Report No. CC-0201). Denver, CO: Education Commission of the States.

Keck, C. (2014). Copying, paraphrasing, and academic writing development: A re-examination of L1 and L2 summarization practices. *Journal of Second Language Writing*, *25*, 4–22. doi:10.1016/j.jslw.2014.05.005

Kitsantas, A., & Zimmerman, B. J. (2009). College students' homework and academic achievement: The mediating role of self-regulatory beliefs. *Metacognition and Learning*, *4*(2), 97–110. doi:10.1007/s11409-008-9028-y

Kiuhara, S. A., O'Neill, R. E., Hawken, L. S., & Graham, S. (2012). The effectiveness of teaching 10th-grade students STOP, AIMS, and DARE for planning and drafting persuasive text. *Exceptional Children*, *78*(3), 335–355. doi:10.1177/001440291207800305

Kuh, G. D., Jankowski, N., Ikenberry, S. O., & Kinzie, J. (2014). *Knowing what students know and can do: The current state of student learning outcomes assessment in U.S. colleges and universities*. Retrieved from National Institute for Learning Outcomes Assessment website: http://www.learningoutcomeassessment.org /documents/2013%20Survey%20Report%20Final.pdf

Kwong, T. E., & Brachman, K. J. (2014). Strategy choice mediates the link between auditory processing and spelling. *PLOS ONE*, *9*(9), e107131. doi:10.1371/journal.pone.0107131

Kwong, T. E., & Varnhagen, C. K. (2005). Strategy development and learning to spell new words: Generalization of a process. *Developmental Psychology*, *41*(1), 148–159. doi:10.1037/0012-1649.41.1.148

Leist, C. W., Woolwine, M. A., & Bays, C. L. (2012). The effects of using a critical thinking scoring rubric to assess undergraduate students' reading skills. *Journal of College Reading and Learning*, *43*(1), 31–58.

Lesaux, N. K., Kieffer, M. J., Kelley, J. G., & Harris, J. R. (2014). Effects of academic vocabulary instruction for linguistically diverse adolescents: Evidence from a randomized field trial. *American Educational Research Journal*, *51*(6), 1159–1194. doi:10.3102/0002831214532165

Liao, H.-A., Edlin, M., & Ferdenzi, A. C. (2014). Persistence at an urban community college: The implications of self-efficacy and motivation. *Community College Journal of Research and Practice*, *38*(7), 595–611. doi:10.1080/10668926.2012.676499

Liu, O. L., Bridgeman, B., & Adler, R. M. (2012). Measuring learning outcomes in higher education: Motivation matters. *Educational Researcher*, *41*(9), 352–362. doi:10.3102/0013189X12459679

Lym, W. L. (2014). Strategies for improving vertical alignment: Implementation of a college-readiness collaborative of secondary and post-secondary educators. *Community College Journal of Research and Practice*, *38*(11), 1053–1056. doi:10.1080/10668926.2013.840689

MacArthur, C. A., & Lembo, L. (2009). Strategy instruction in writing for adult literacy learners. *Reading and Writing: An Interdisciplinary Journal*, *22*(9), 1021–1039. doi:10.1007/s11145-008-9142-x

MacArthur, C. A., & Philippakos, Z. (2010). Instruction in a strategy for compare–contrast writing. *Exceptional Children*, *76*(4), 438–456. doi:10.1177/001440291007600404

MacArthur, C. A., & Philippakos, Z. A. (2013). Self-regulated strategy instruction in developmental writing: A design research project. *Community College Review*, *41*(2), 176–195. doi:10.1177/0091552113484580

MacArthur, C. A., Philippakos, Z. A., & Graham, S. (2015). A multicomponent measure of writing motivation with basic college writers. *Learning Disability Quarterly*. Advance online publication. doi:10.1177/0731948715583115

MacArthur, C. A., Philippakos, Z. A., & Ianetta, M. (2015). Self-regulated strategy instruction in college developmental writing. *Journal of Educational Psychology*, *107*(3), 855–867. doi:10.1037/edu0000011

Macaruso, P., & Shankweiler, D. (2010). Expanding the simple view of reading in accounting for reading skills in community college students. *Reading Psychology*, *31*(5), 454–471. doi:10.1080/02702710903241363

Magliano, J. P., Trabasso, T., & Graesser, A. C. (1999). Strategic processing during comprehension. *Journal of Educational Psychology*, *91*(4), 615–629. doi:10.1037/0022-0663.91.4.615

Martin, S. D., & Shapiro, E. S. (2011). Examining the accuracy of teachers' judgments of DIBELS performance. *Psychology in the Schools*, *48*(4), 343–356. doi:10.1002/pits.20558

Martinez, C. T., Kock, N., & Cass, J. (2011). Pain and pleasure in short essay writing: Factors predicting university students' writing anxiety and writing self-efficacy. *Journal of Adolescent and Adult Literacy*, *54*(5), 351–360. doi:10.1598/JAAL.54.5.5

Martorell, P., & McFarlin, I., Jr. (2011). Help or hindrance? The effects of college remediation on academic and labor market outcomes. *Review of Economics and Statistics*, *93*(2), 436–454. doi:10.1162/REST_a_00098

Mateos, M., Martín, E., Villalón, R., & Luna, M. (2008). Reading and writing to learn in secondary education: Online processing activity and written products in summarizing and synthesizing tasks. *Reading and Writing: An Interdisciplinary Journal*, *21*(7), 675–697. doi:10.1007/s11145-007-9086-6

McAlexander, P. J. (2003). From personal to text-based writing: The use of readings in developmental composition. *Research and Teaching in Developmental Education*, *19*(2), 5–16.

McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written Communication*, *27*(1), 57–86. doi:10.1177/0741088309351547

Melguizo, T., Bos, J., & Prather, G. (2011). Is developmental education helping community college students persist? A critical review of the literature. *American Behavioral Scientist*, *55*(2), 173–184. doi:10.1177/0002764210381873

Mellard, D. F., Krieshok, T., Fall, E., & Woods, K. (2013). Dispositional factors affecting motivation during learning in adult basic and secondary education programs. *Reading and Writing: An Interdisciplinary Journal*, *26*(4), 515–538. doi:10.1007/s11145-012-9413-4

Merchie, E., & Van Keer, H. (2014). Learning from text in late elementary education. Comparing think-aloud protocols with self-reports. *Procedia - Social and Behavioral Sciences*, *112*, 489–496. doi:10.1016/j.sbspro.2014.01.1193

Merriam, S. B., & Bierema, L. L. (2013). *Adult learning: Linking theory and practice*. New York, NY: Jossey-Bass/Wiley.

Mongillo, G., & Wilder, H. (2012). An examination of at-risk college freshmen's expository literacy skills using interactive online writing activities. *Journal of College Reading and Learning*, *42*(2), 27–50. doi:10.1080/10790195.2012.10850353

Moore, N. S., & MacArthur, C. A. (2012). The effects of being a reader and of observing readers on fifth-grade students' argumentative writing and revising. *Reading and Writing: An Interdisciplinary Journal*, *25*(6), 1449–1478. doi:10.1007/s11145-011-9327-6

Nash-Ditzel, S. (2010). Metacognitive reading strategies can improve self-regulation. *Journal of College Reading and Learning*, *40*(2), 45–63. doi:10.1080/10790195.2010.10850330

National Center for Education Statistics. (2012). *The nation's report card: Writing 2011* (NCES 2012–470). Retrieved from National Center for Education Statistics website: http://nces.ed.gov/nationsreportcard/pdf/main2011/2012470.pdf

National Center for Education Statistics. (2014). *The nation's report card: Are the nation's 12th-graders making progress in mathematics and reading?* (NCES 2014–087). Retrieved from National Center for Education Statistics website: http://nces.ed.gov/nationsreportcard/subject/publications/main2013/pdf/2014087.pdf

National Center on Education and the Economy. (2013). *What does it really mean to be college and work ready? The mathematics and English literacy required of first year community college students.* Retrieved from NCEE website: http://www.ncee.org/college-and-work-ready/

National Governors' Association Center for Best Practices, & Council of Chief State School Officers. (2010). *Common Core State Standards: English language arts and literacy in history/social studies, science, and technical subjects.* Washington, DC: Author.

Nelson, N. W., & Van Meter, A. M. (2007). Measuring written language ability in narrative samples. *Reading and Writing Quarterly*, *23*(3), 287–309. doi:10.1080/10573560701277807

Newell, G. E., Beach, R., Smith, J., & VanDerHeide, J. (2011). Teaching and learning argumentative reading and writing: A review of research. *Reading Research Quarterly*, *46*(3), 273–304. doi:10.1598/RRQ.46.3.4

Nunley, C., Bers, T., & Manning, T. (2011). *Learning outcomes assessment in community colleges* (Occasional Paper No. 10). Retrieved from National Institute for Learning Outcomes Assessment website: http://www.learningoutcomesassessment.org/documents/CommunityCollege.pdf

Olinghouse, N. G. (2008). Student- and instruction-level predictors of narrative writing in third-grade students. *Reading and Writing: An Interdisciplinary Journal*, *21*(1–2), 3–26. doi:10.1007/s11145-007-9062-1

Olinghouse, N. G., & Wilson, J. (2013). The relationship between vocabulary and writing quality in three genres. *Reading and Writing: An Interdisciplinary Journal*, *26*(1), 45–65. doi:10.1007/s11145-012-9392-5

O'Neill, P., Adler-Kassner, L., Fleischer, C., & Hall, A.-M. (2012). Creating the framework for success in postsecondary writing. *College English*, *74*(6), 520–533.

Pajares, F., & Valiante, G. (2006). Self-efficacy beliefs and motivation in writing development. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 158–170). New York, NY: Guilford Press.

Parsad, B., & Lewis, L. (2003). *Remedial education at degree-granting postsecondary institutions in fall 2000: Statistical analysis report* (NCES 2004-010) Washington D.C.: U.S. Department of Education, National Center for Education Statistics.

Paulson, E. J. (2014). Analogical processes and college developmental reading. *Journal of Developmental Education*, *37*(3), 2–13.

Perin, D. (2011). Facilitating student learning through contextualization: A review of the evidence. *Community College Review*, *39*(3), 268–295. doi:10.1177/0091552111416227

Perin, D., Bork, R. H., Peverly, S. T., & Mason, L. H. (2013). A contextualized curricular supplement for developmental reading and writing. *Journal of College Reading and Learning*, *43*(2), 8–38. doi:10.1080/10790195.2013.10850365

Perin, D., & Charron, K. (2006). "Lights just click on every day." In T. Bailey & V. S. Morest (Eds.), *Defending the community college equity agenda* (pp. 155–194). Baltimore, MD: Johns Hopkins University Press.

Perin, D., Keselman, A., & Monopoli, M. (2003). The academic writing of community college remedial students: Text and learner variables. *Higher Education*, *45*(1), 19–42. doi:10.1023/A:1021237532056

Pimentel, S. (2013). *College and career readiness standards for adult education*. Retrieved from Literacy Information and Communication System website: https://lincs.ed.gov/publications/pdf/CCRStandardsAdultEd.pdf

Plakans, L. (2008). Comparing composing processes in writing-only and reading-to-write test tasks. *Assessing Writing*, *13*(2), 111–129. doi:10.1016/j.asw.2008.07.001

Porter, A. C., & Polikoff, M. S. (2012). Measuring academic readiness for college. *Educational Policy*, *26*(3), 394–417. doi:10.1177/0895904811400410

Proctor, C. P., Daley, S., Louick, R., Leidera, C. M., & Gardner, G. L. (2014). How motivation and engagement predict reading comprehension among native English-speaking and English-learning middle school students with disabilities in a remedial reading curriculum. *Learning and Individual Differences*, *36*, 76–83. doi:10.1016/j.lindif.2014.10.014

Puranik, C. S., Lombardino, L. J., & Altmann, L. J. P. (2008). Assessing the microstructure of written language using a retelling paradigm. *American Journal of Speech-Language Pathology*, *17*(2), 107–120. doi:10.1044/1058-0360(2008/012)

Retelsdorf, J., Köller, O., & Möller, J. (2011). On the effects of motivation on reading performance growth in secondary school. *Learning and Instruction*, *21*(4), 550–559. doi:10.1016/j.learninstruc.2010.11.001

Reynolds, G. A., & Perin, D. (2009). A comparison of text-structure and self-regulated writing strategies for composing from sources by middle school students. *Reading Psychology*, *30*(3), 265–300. doi:10.1080/02702710802411547

Ritchey, K. D., Silverman, R. D., Schatschneider, C., & Speece, D. L. (2015). Prediction and stability of reading problems in middle childhood. *Journal of Learning Disabilities*, *48*(3), 298–309. doi:10.1177/0022219413498116

Robinson, D. H., & Kiewra, K. A. (1995). Visual argument: Graphic organizers are superior to outlines in improving learning from text. *Journal of Educational Psychology*, *87*(3), 455–467. doi:10.1037/0022-0663.87.3.455

Salvia, J., Ysseldyke, J. E., & Bolt, S. (2013). *Assessment in special and inclusive education* (12th ed.). Belmont, CA: Wadsworth/Cenage Learning.

Schrank, F. A., McGrew, K. S., & Woodcock, R. W. (2001). *Woodcock-Johnson III* (Assessment Service Bulletin No. 2). Itsaca, IL: Riverside Publishing.

Scott-Clayton, J. (2012). *Do high-stakes placement exams predict college success?* (CCRC Working Paper No. 41). New York, NY: Columbia University, Teachers College, Community College Research Center.

Shanahan, C., Shanahan, T., & Misischia, C. (2011). Analysis of expert readers in three disciplines: History, mathematics, and chemistry. *Journal of Literacy Research*, *43*(4), 393–429. doi:10.1177/1086296X11424071

Simpson, M. L. (1986). PORPE: A writing strategy for studying and learning in the content areas. *Journal of Reading*, *29*(5), 407–414.

Sparks, D., & Malkus, N. (2013). *First-year undergraduate remedial coursetaking: 1999–2000, 2003–04, 2007–08* (NCES 2013-013). Retrieved from National Center for Education Statistics website: http://nces.ed.gov/pubs2013/2013013.pdf

Speece, D. L., Ritchey, K. D., Silverman, R., Schatschneider, C., Walker, C. Y., & Andrusik, K. N. (2010). Identifying children in middle childhood who are at risk for reading problems. *School Psychology Review*, *39*(2), 258–276.

Steffler, D. J., Varnhagen, C. K., Friesen, C. K., & Treiman, R. (1998). There's more to children's spelling than the errors they make: Strategic and automatic processes for one-syllable words. *Journal of Educational Psychology*, *90*(3), 492–505. doi:10.1037/0022-0663.90.3.492

Strømsø, H. I., Bråten, I., Britt, M. A., & Ferguson, L. E. (2013). Spontaneous sourcing among students reading multiple documents. *Cognition and Instruction*, *31*(2), 176–203. doi:10.1080/07370008.2013.769994

Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, *104*(3), 743–762. doi:10.1037/a0027627

Theurer, J. L. (2011). Does accuracy make a difference? Examining the miscues of proficient and less than proficient adult readers. *Literacy Research and Instruction*, *50*(3), 173–182. doi:10.1080/19388071003797449

Tillema, M., van den Bergh, H., Rijlaarsdam, G., & Sanders, T. (2011). Relating self-reports of writing behaviour and online task execution using a temporal model. *Metacognition and Learning*, *6*(3), 229–253. doi:10.1007/s11409-011-9072-x

Torraco, R. J. (2014). Remedial education: An area in need of scholar-practitioner collaboration. *Community College Journal of Research and Practice*, *38*(12), 1198–1202. doi:10.1080/10668926.2014.899527

Troia, G. A., Harbaugh, A. G., Shankland, R. K., Wolbers, K. A., & Lawrence, A. M. (2013). Relationships between writing motivation, writing activity, and writing performance: Effects of grade, sex, and ability. *Reading and Writing: An Interdisciplinary Journal*, *26*(1), 17–44. doi:10.1007/s11145-012-9379-2

Uccelli, P., Dobbs, C. L., & Scott, J. (2013). Mastering academic language: Organization and stance in the persuasive writing of high school students. *Written Communication*, *30*(1), 36–62. doi:10.1177/0741088312469013

Wang, D. (2009). Factors affecting the comprehension of global and local main idea. *Journal of College Reading and Learning*, *39*(2), 34–52. doi:10.1080/10790195.2009.10850317

Westby, C., Culatta, B., Lawrence, B., & Hall-Kenyon, K. (2010). Summarizing expository texts. *Topics in Language Disorders*, *30*(4), 275–287. doi:10.1097/TLD.0b013e3181ff5a88

Williamson, G. L. (2008). A text readability continuum for postsecondary readiness. *Journal of Advanced Academics*, *19*(4), 602–632. doi:10.4219/jaa-2008-832

Wolfe, C. R. (2011). Argumentation across the curriculum. *Written Communication*, *28*(2), 193–219. doi:10.1177/0741088311399236

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement and Tests of Cognitive Abilities.* Itasca, IL: Riverside Publishing.

Yancey, K. B. (2009). The literacy demands of entering the university. In L. Christenbury, R. Bomer & P. Smagorinsky (Eds.), *Handbook of adolescent literacy research* (pp. 256–270). New York, NY: Guilford Press.

Appendix A

Examples of Student Writing

Scores shown are for analytic essay quality and holistic persuasive quality.

Essay prompt: Is bad behavior in teens due to stress? What is your opinion? Pretend you are trying to persuade a friend to agree with you.

Student 1

Summary score = 9

Generation Y, a generation of younger and more faster group of people. Working harder and more smarter than the past generation. Give more benefits to the working job because they know about computers and how people may now think. They tend to be around the age of 20 and can multitasks. In the workplace their having no problems other than the conflict of having to tell someone older what to do.

Essay score = 3

Bad behavior from teens is due to be from stress. A research was shown on how stress in teens could lead to the way they might act. 26 percent of teens reported snapping with classmates from stress. I feel as though stress can be a big hold on teens actions. While it might not be the main reason, most teens often become overwhelmed with problems. A highly number of adults think that students aren't stressing or may not have problems, but their wrong. Being a teen myself, I know how it feels to stress out from school, work, and friends. While some people get the privilege of not having to get a job others teens don't. Often stress levels can as well come from home and as well as school. There are many different ways a teen might become stressed, but they all react differently. While some might ignore it others might go in a deep depression or even become suicidal. Not every over stressed teen would go crazy but most of the time it will have an impact and we need to change that.

Student 2

Summary score = 12

This article talks about the younger generation entering the work-force ready to make an impact. The college graduates are entering the work-force at a major demographic change. Companies in the United States are facing a aging workforce. Generation Y are not only young but high performance and high maintenance. There the generation that can multitask and change companies to make them better. Generation Y is also stirring up conflict with the old employees who have worked at a company for over 10 years. these older employees don't want to have to report to young students. Overall, this article talks about how generation y is taking over the workplace.

Essay score = 2

I believe that bad behavior in teens can come from some stress but not all of it. Behavior in teens is how they act and you cannot blame stress for kids going out and making bad decisions. Those kids make those decisions because they want to. Although some stress can make teens want to relax by drinking and smoking. I believe that only a small percent of stress leads to students making bad decisions. You as a human being are responsible for your own actions if you don't want to make a bad decision that you want. When I'm stressed I tend to workout more and it makes me feel better. I think students should take there stress out on other stuff instead of making wrong decisions. Teens and adults should channel there stress into something else besides not getting anything done, making wrong decisions, and lashing out at other people.

Student 3

Summary score = 11

The Generation Y's have arrived at workplace with a new attitude. They are young, smart, and brash. They are also computer savvy. There are tensions between generations because both Generations are not appreciating the other generation. It is older workers are now reporting to someone young enough to be their child. The Generation Y have been pampered nurtured and have been involved with a lot of activities since they were toddlers, so they know their own worth. The newest members of the workforce are smart where saving money. They want jobs that are flexible, telecommuting, the option to go part time, or leave Temp. When have kids. They do not have plans to stay in the job or even the career for long. They are very good at multitasking and they do not like to stay on 1 thing for long. They believe in themselves and their value that they are not shy about changing the company. The Generation Y total tech savvy compared to older generation b/c the older generation would expect to get a call about coming in for a meeting but the generation Y are more adept at having a meeting on the computer. The gen Y don't get respect deserved b/c of their age.

Essay score = 2

When teens exhibit bad behavior it is due to stress sometimes but not always. Teens today are more stressed than before but I don't believe that it is the case of them acting out. I think that teens use being stressed out as an excuse so they won't get in trouble. However some teens are really stressed out. They are depressed, anxious, anger, have withdrawn & have ongoing irritability. They copy their parents actions when their parents are stressed. My opinion is that kids today will say that they are stressed to explain their behavior and then copy their parents actions when stressed. I don't believe that stress is the factor of bad behavior. I think something else is going on and they use being stressed as an excuse. It's bad because of the kids that do this it makes it harder for the kids that are actually stressed because the doctors don't know whether to believe them or not.

Student 4

Summary score = 5

This article about Generation Y is so right because it more younger people in management positions than older. More older people are have to answer to a younger person. This is only because Generation Y have a huge thing for technology. Technology is running the world so it put younger people at an advance. Older people don't have experience with computers like Generation Y.

Essay score = 3

Bad behavior can and also cannot be cause by bad behavior. in my own opinion, I think that stress can cause bad behavior such as snapping, lashing out at family, and trying drugs. I also think that it not the cause for all bad behaviors. I just don't think stress cause you to go steal from somebody or somewhere. I also think that stress don't cause you to go kill nobody or cause them to commit suicide. No I won't believe that stress causes people to do a lot of things to put them in jail.

> Pretending:
> Me: Hay Kim
> Kim: Hay
> Me: Can you believe that teacher saying stress cause a lot of teen behaviors.
> Kim: Yes I can believe that
> Me: What?
> Kim: Yes I believe it do
> Me: Well listen to this! Do you think stress causes a teen to go kill someone over what going they with? Do you think stress causes you or them to steal? What if they did it to you?
> Kim: You have a point because stress don't cause you to pull a trigger or steal from anyone.
> Me: Thank you! My point exactly.

Student 5

Summary score = 4

Generation Y is talking about a workforce. The things the talk about in the workforce are the ages of the workers. They focus on how well people work and perform task. Last it focus on younger people more than adults.

Essay score = 2

I don't think stress is always cause by bad behaviors. Some bad behavior can make a teen stress. But bad behaviors is not what really can cause stress. For example for some teens if they get a bad grade on a test that can cause them to stress. Another example is if the teen is having a hard time at home with family it can cause them to stress. Last example can be if the teen is fighting with there best friend it can make them stress. So we can see

by this article stress can be cause by a lot of different thing. Don't think stress is only cause by bad behavior when everything basically can cause teens to stress. There in the mid-adulthood so there will be a lot of stress. Last but not least parents to help there teen out if they are stressing.

Student 6

Summary score = 7

[Title of article] written by [name of author] is an article about how our generation today is different from the old generation. How children today are use to getting constant approval and feedback from teachers, and parents when they enter the work force they feel lost when there boss doesn't give them constant approval. This article also discusses how it can be harder for an older generation to work with a newer generation. They have completely different mindsets. The newer generation can work and play on phone but still get there job done. For an older generation you are not suppose to do that. Its unprofessional. In the end of this article it states that there is advantages to both younger and older generation.

Essay score = 4

In my opinion I do not think that to some extent teens bad behavior is due to stress. Not everyone might agree but teens tend to have a lot of stress of completing assignments for school, work, chores and other things. Now this doesn't sound like a lot but when your a teen you stress more about silly, unimportant things then you do when you are an adult. Teens stress about being perfect, doing things perfectly which can cause high levels of stress. Clinical psychologist Jonathan Abramowitz makes a good point that there is no sure way of telling if a teens bad behavior is from stress but I believe some adolescents especially recently have increased with having depression which causes teens to be stressed. I know from my own experience that when I was a adolescent I strived for perfection which put a lot of pressure on me and caused me to act out and feel depressed. Now that I am older I realize there is no exact reason for me to act that way even if I felt pressured or stressed. [Name], a psychologist in [location], thinks that parents can only relive there adolescents stress to an extent. In my opinion I think a parent helping can create more stress for a teen because they are focused on thinking there parents believe something is wrong with them. Either way it's hard to tell how you can help an adolescent feel less stressed and if there bad behaviors is from stress.

Student 7

Summary score = 9

Young verse older employees is becoming more of a conflict in the workplace now more than the past decades. Conflict simply arrives when there is a person that has been on the job for several years that know about the company and how it operates, but then a younger person is hired in a position higher that see and understand a better way or much easier way that the company can operate at the same place or a better place. The conflict occurs when the dress attire is different, the person set to be the overseer is young, and

the attitude of the co-workers (what does he know) come in. But this is a common response to any situation in life about change. People have been doing things a certain way for so long change scares them. People fear what they don't understand. If the company is still gaining and not losing in any way change can be good beside the younger generation is our future.

Essay score = 2

Stress can be good or bad, its not what you say bad behavior it normal for people to stress. It's the amount of stress you hold that makes it healthy or unhealthy. I think if students learn to control their stress and what they think about that causes them stress that stress can easily be brought under control. I come up with a saying "it is what it is" and, that saying means to me, I can't worry about things I can't change so, there for my stress level stays the same. As research mention some teens merely used this as an excuse, way to get attention. We've all done it before, took the smallest problems and made it into a crisis.

Student 8

Summary score = 11

Evolution has taken its place, and Gen. Y has begun starting in the workforce. There are younger adults that are now as qualified to be working next to adults that are twice their age. Gen Y are higher maintenance, but can multitask with todays work tools. Also there are cases where Older adults not have to report to someone as young as their child. All in all Gen Y has learned from previous generation and they are a new breed in the workforce.

Essay score = 2

All over the country, researchers have concerns whether bad behavior is linked with stress. I feel this link not to be true because I have had a very stressful life and I'm fine, I have personally experienced teens using every excuse in the book to take less responsibility, and stress prepares teens for the future. First my life has been very stressful since the passing of my father when I was 3 years old. My mother and I were basically homeless because of financial reasons. Also, growing up and going to school I was teased for my A.D.D. Now I live in North Carolina, and my family lives in Texas. These are all the reasons in the world to be stressed, yet I am not out there robbing banks. Next, I have personally experienced many teens use all kinds of stress related excuses, or any excuse for that matter, to take less responsibility for their actions. For example, John Doe got in trouble for bullying, and he blamed everything on stress and A.D.D. I know from experience that he was full of bull because I have A.D.D. and I am no bully. Also Mr. Doe blamed stress for not doing his homework. In my opinion there is not enough stress in the world that would keep you.

Appendix B

Student Self-Efficacy Ratings and Teacher Judgments

Appendix Table B.1
Mean Self-Efficacy Ratings and Teacher Judgments, Total Sample (*N* = 211)

Student Question	Student Mean (*SD*)	Teacher Question	Teacher Mean (*SD*)	Correlation
1. I can read short newspaper articles and understand them well.	86 (14.9)	1. The student can read passages such as the attached with the level of understanding expected at an introductory college level.	75.2 (18.1)	.042 *n* = 207
2. I can read the articles carefully and form my own opinion about the issues discussed.	86.4 (15.9)	2. The student can read the passages carefully and think critically about the ideas discussed.	72.9 (19.3)	.123 *n* = 204
3. I can figure out how information in a newspaper article might be useful.	82.7 (16)	3. The student can figure out how information in passages such as the attached might be useful.	74.6 (16.2)	.156* *n* = 198
4. I can write a good summary of a short article from a newspaper.	78 (18.2)	4. The student can write good summaries based on reading passages, like the ones attached, at the college level.	68.3 (19.1)	.164* *n* = 199
5. I can write a summary of a newspaper article that includes only the most important information.	79.3 (18.2)	5. The student can write a summary that includes only the most important information from a passage.	69.2 (17.5)	.209** *n* = 198
6. I can write an essay that expresses my opinion clearly.	83.1 (17.9)	6. The student can write a persuasive essay that expresses his or her opinion clearly.	68.1 (17.5)	.248** *n* = 195
7. In my essay I can persuade someone to agree with me on my opinion.	75.2 (18.3)	7. The student can write a persuasive essay that effectively convinces a reader of his or her position on a topic.	65.2 (19)	.125 *n* = 194
8. If I write a summary or essay based on something I have read, I can express the information from the reading accurately.	79.2 (17.3)	8. If the student can write a summary or essay based on a reading passage, he or she can present information from the passage accurately.	70 (18.6)	.173* *n* = 198
9. I can write a summary or essay in my own words, without copying directly from a reading passage.	81.2 (19.1)	9. The student can write a summary or essay in his or her own words, without copying directly from a reading passage.	71.7 (17.8)	.200** *n* = 199
10. I can write a summary or essay using correct grammar and spelling.	72.2 (21.6)	10. The student can write a summary or essay using correct grammar and spelling.	63.6 (19.2)	.305** *n* = 197

Student Question	Student Mean (*SD*)	Teacher Question	Teacher Mean (*SD*)	Correlation
11. I can write a summary or essay that is the right length—not too long, not too short.	78.2 (18.6)	11. The student can write a summary or essay of an appropriate length for college.	74.1 (18.8)	.242** *n* = 200
12. I can write a summary or essay using appropriate academic vocabulary.	73.9 (20.7)	12. The student can write a summary or essay using appropriate academic vocabulary.	66.5 (19.9)	.33** *n* = 201
13. I can revise my summary or essay to make sure what I've written is accurate and clear.	79.7 (19.1)	13. The student can revise his or her summary or essay to ensure accuracy and clarity.	66.9 (19.7)	.114 *n* = 201
14. I will be able to understand the reading in the courses I take after I pass the English or reading class I'm in now.	83.9 (21.5)	14. The student will be able to read with adequate understanding in most future college classes.	75.6 (15.8)	.187* *n* = 203
15. I will be able to write well in the courses I take after I pass the English or reading class I'm in now.	85.3 (15.1)	15. The student will be able to handle the writing required in most future college-level classes.	70.5 (18.6)	.216** *n* = 196
16. When reading or writing assignments are hard, I keep going and finish the assignment.	86.1 (18.4)	16. The student is able to work with sustained effort to finish reading and writing tasks that are difficult for him or her.	76.4 (18.7)	.140 *n* = 194

* $p < .05$. ** $p < .01$.